Searches for Pair Production of Heavy Vector-like Quarks in pp Collisions at $\sqrt{s}$ =13 TeV with the ATLAS Detector

by

Daniel Edison Marley

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2017

Doctoral Committee:

Assistant Professor Thomas A. Schwarz, Chair
Professor Dante E. Amidei
Professor Gordon L. Kane
Professor Sara A. Pozzi
Professor Jianming Qian

Daniel Edison Marley

demarley@umich.edu

ORCID iD: 0000-0002-6290-078X

To my parents.

ACKNOWLEDGEMENTS

The work described here would not have been possible without substantial support and encouragement from so many family, friends, and colleagues. I cannot express enough gratitude to the people that have played a role in my journey thus far, but I will try nonetheless.

Tom Schwarz has been an exceptional advisor. His direction and support of my research efforts helped me to find ways to succeed in the ATLAS collaboration. Tom has been an example of how to have a healthy work-life balance while maintaining a high level of productivity in the collaboration. The freedom to explore my own ideas (and sometimes preventing me from going off in *too* many directions) has been invaluable in my development as a physicist.

My committee members Dan Amidei, Gordy Kane, Sara Pozzi, and Jianming Qian have been very supportive through this process and their feedback has been invaluable.

I'm very grateful to my undergraduate advisor Paul Huffman for offering me my first research opportunity, despite my lack of experience, and continually providing opportunities to develop as an experimental physicist. I'm also grateful to the entire NC State faculty for providing me with an strong education that has positioned me well for pursuing experimental physics research.

My interest in scientific research would not be what it is today without the education I received from my high school chemistry and physics teacher, Glenn Powell. Mr. Powell was the first teacher who challenged and pushed me to think in new and creative ways. His encouragement to pursue physics has enabled me to find a career that is both rewarding and challenging every day.

I've been lucky enough to work with many intelligent and enthusiastic people in the ATLAS collaboration, from whom I've learned a tremendous amount about particle and collider physics. Francesco Rubbo and Clement Helsens were incredibly helpful and patient while teaching me about top quarks and unfolding. Marcel Vos and Joe Haley provided a wealth of information and support. I learned an immense amount of physics from our many discussions.

I would also like to thank Ece Akilli, Olaf Nackenhorst, Madhuranga Thilakasiri, and David Jamin for contributing to the VLQ 0-lepton analysis with enthusiasm and hard work.

While at UM I've worked with very bright and enthusiastic undergraduate students: Bennett Magy, Ronak Mehta, Adrian Sanchez, and Aaron White. Your unique perspectives and hard work were much appreciated. I take credit for all of your future success.

I want to thank the Fondue crew: Allison McCarn, Hao Liu, Rachel Hyneman, and Garrett Merz. Our lunch breaks and engaging discussions have been a rewarding part of sharing an office and collaborating with you all.

Being a member of the UM ATLAS group has been rewarding in many ways and I'm very thankful for having worked in this group. The group has always fostered an environment of creativity and encouraged me to pursue innovative research ideas.

Many of the people I met at Michigan have made graduate school a wonderful experience: Joe Osborn for foot golfing, college sports debates, and research walks; Uri Israel for countless discussions about machine learning; Andre Thompson, Eric Warner & the rest of the Rackhammers for the welcome distraction of flag football (and multiple intramural championships!).

My family has provided unwavering support and encouragement. I want to especially thank my parents, Larry and Marty, for instilling an appreciation for education, always believing in me, and pushing me to do my best. Regardless of what is going on, my parents were always eager to hear about my adventures at Michigan. I also want to thank my siblings (in birth order, not necessarily who I like the most right now): Kaitlin, Josh, & Luke, Kristin & Andrew, Robert, and David for their continued support. Without them I wouldn't be nearly as competitive or driven to succeed (you all make the holidays more enjoyable, too).

My wonderful wife Jenny deserves all the thanks in the world. I can't begin to describe the level of support she has provided. All of our adventures together have been filled with endless amounts of excitement and made grad school that much more enjoyable. I'm incredibly proud of your own accomplishments at Michigan. Moving from Raleigh to Ann Arbor with you was one of the best decisions of my life and I cannot wait to see what the next chapter has in store. My gratitude towards Jenny also extends to her family. They have been remarkably supportive of us and are always there to provide us with anything we need.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

A summary of two orthogonal and complementary searches for pair production of heavy vector-like quarks is presented. The first analysis considers final states with 1-lepton and hadronic jets that can be used to reconstruct the vector-like quark system after identifying the boosted decay of W bosons and building the neutrino from the missing transverse momentum. Control and signal regions constrain the uncertainties associated with $t\bar{t}$ production that most negatively impact this analysis. The second analysis searches in the fully-hadronic final state where no leptons are considered. An explicit veto on the lepton number ensures orthogonality between the analyses. The analysis uses a deep neural network to reconstruct the decays of heavy objects and a data-driven technique to estimate the dominant multi-jet background. Data were recorded by the ATLAS experiment in $\sqrt{s}$=13 TeV proton-proton collisions delivered by the Large Hadron Collider in 2015 and 2016. No excess above the Standard Model background is observed in either analysis. The results from the 1-lepton analysis are interpreted assuming vector-like quarks decay $T\bar{T} \to Wb + X$. The 0-lepton analysis assumes both types of pair produced vector-like quarks $T\bar{T}$ and $B\bar{B}$ that decay to $T \to Wb, Ht, Zt$ and $B \to Wt, Hb, Zb$. The 1-lepton analysis observes a 95% CL lower limit on the T mass of 1090 (810) GeV assuming the scenario BR($T \to Wb$) = 1 ($SU(2)$ singlet). The strongest observed 95% CL lower limits in the 0-lepton analysis are 850 GeV and 903 GeV in the BR($T \to Ht$) = 1 and BR($B \to Hb$) = 1 decays, respectively.

CHAPTER I

Introduction

At the European Organization for Nuclear Research (Conseil Européen pour la Recherche Nucléaire, CERN), the Large Hadron Collider is colliding the most energetic beams of particles to recreate conditions just after the big bang. Two beams of protons accelerate in opposite directions to nearly the speed of light and collide at four separate points along the collider ring. **A T**oroidal **L**HC **A**pparatu**S** (ATLAS) is located at one of these intersections to capture three-dimensional images of these collisions using a variety of particle detector technologies. Nearly 3000 physicists collaborate on the ATLAS experiment to operate, maintain, and develop the detector technologies while analyzing the data collected.

The theoretical framework of particle physics, known as the Standard Model, describes all known particles and their interactions. However, a few open questions remain in particle physics, unanswered by the Standard Model, such as what is the nature of dark matter and why the Higgs boson has a relatively small mass despite quadratic divergences in the mass calculation.

In 2015, the Large Hadron Collider began colliding proton beams at a center-of-mass energy of 13 TeV, compared to 8 TeV in 2012. The jump in energy renewed interest in searches for new physics given significant increases in production cross sections. Furthermore, the tremendous amount of data provided by the LHC and collected by ATLAS in 2015 and 2016 improves the sensitivity to new physics for analyses.

A subset of the ATLAS research program is dedicated to uncovering new physics. Such efforts include the searches for vector-like quarks. Vector-like quarks are the simplest colored fermion extra generation still allowed by experimental results and they can naturally regulate the Higgs mass. Additionally, vector-like quarks fit into a wide variety of theoretical models, further motivating their searches. The ATLAS and Compact Muon Solenoid (CMS) collaborations have performed a wide variety of searches for vector-like quarks using a multitude of final states and assuming both pair and single production [2–16].

This dissertation describes two searches for vector-like quarks performed with the ATLAS detector at the Large Hadron Collider. Chapter II introduces the theoretical framework of particle physics and motivations to searchi for new physics. The experimental apparatuses, both ATLAS and the Large Hadron Collider, are described in Chapter III. Experimental methods for translating detector data into physics objects are discussed in Chapter IV. The analyses and results are detailed in Chapter V. Chapter VI summarizes the research efforts in this dissertation.

CHAPTER II

Theoretical Framework

The Standard Model (SM) of particle physics describes nature through elementary particles and their interactions. The SM has consistently agreed with a wide variety of experimental results, including the physics observed at particle colliders. The theory was experimentally validated with the discovery of the Higgs bosons in 2012 [17, 18]. There are observed phenomena unexplained by the SM, see Sect. 2.2. Much of the effort at the Large Hadron Collider is dedicated to understanding the sources of these discrepancies through precision measurements of SM properties and direct searches for physics beyond the SM (BSM). This section reviews the formulation of the SM followed by examples of limitations of the theory and a discussion on vector-like quark phenomenology.

2.1 The Standard Model of Particle Physics

At distances near and smaller than an atom, physical processes are best described using quantum mechanics. As the energy scale increases, resolving quantum mechanics with special relativity results in *quantum field theory* (QFT). The SM is a specific formulation of QFT [19–22]. The description of physics in QFT relies on fields and their properties in space-time. Fundamental particles are represented as excitations of a field, ensuring the existence of distinct, but identical particles for each field. Particles in the SM with half-integer spin that obey Fermi-Dirac statistics are known as *fermions*. These particles all interact via force carriers referred to as *bosons*. The bosons in the SM have integer spin and obey Bose-Einstein statistics. The dynamics of the SM are described using the Lagrangian formalism with the principle of least action, $\delta S = 0$, where $S = \int \mathcal{L}(x) d^4x$. The Lagrangian density, $\mathcal{L}$ (hereafter referred to as simply the Lagrangian), contains all of the information about the system: kinematics, couplings, and symmetries.

The three known internal symmetries of the SM are best described using the language of group theory: $U(1)$, $SU(2)$, and $SU(3)$. The full set of symmetries, singularly referred to as the SM gauge group $G_{\rm SM}$, can be written as

$$G_{\rm SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

where the subscripts refer to the charge associated with that group (Y = hypercharge, L = left-handed weak isospin, C = color charge). All known interactions except for gravity are included in $G_{\rm SM}$. The gravitational force is significantly weaker than the other fundamental forces for the processes and scales encountered in this dissertation.

The incredibly rich phenomenology of the SM emerges from requiring the Lagrangian to be invariant with respect to *local* transformations of the fields. As a result, massless spin-1 bosons ($U(1)$: 1 boson, $SU(n)$: n^2-1 bosons) emerge. There are 8 mediators G_i^μ, $i = 1, \ldots, 8$, for the $SU(3)_C$ group symmetry, 1 mediator B^μ for the $U(1)_Y$ symmetry, and 3 mediators W_i^μ, $i = 1, 2, 3$, responsible for the $SU(2)_L$ symmetry.

The $SU(3)$ component of the SM gauge group describes quantum chromodynamics (QCD) where the eight mediators G_i^μ are known as the gluons. The gluons are massless particles that can self-interact because they also carry color charge.

The second two components of $G_{\rm SM}$, $SU(2)_L \otimes U(1)_Y$, represent the electroweak theory. This is a unified theory of the electromagnetic and weak interactions. The bosons B^μ and W_i^μ can be re-written to produce the experimentally observable $W^\pm/Z$ bosons and photon (γ):

$$\begin{aligned} W^\pm &= \big(- W_1^\mu \pm i W_2^\mu \big)/\sqrt{2} \\ \begin{pmatrix} Z \\ \gamma \end{pmatrix} &= \begin{pmatrix} -\sin\theta_W \cos\theta_W \\ \cos\theta_W \sin\theta_W \end{pmatrix} \begin{pmatrix} W_3^\mu \\ B^\mu \end{pmatrix} \end{aligned} \tag{2.2}$$

where θ_W is the weak mixing angle, given by the ratio of W and Z masses: $\cos\theta_W = m_W/m_Z$. The electric charge is related to the electroweak quantum numbers via $Q = T_3 + Y/2$, where T_3 is the $SU(2)_L$ charge (the third-component of the weak isospin T). In this configuration, the Z and photon are involved in neutral current interactions while the $W^\pm$ bosons are responsible for charged current interactions. The photon couples to all electrically charged particles and the $W^\pm/Z$ couple to each other and all fermions.

There are three generations of fermions in the SM, represented in Fig. 2.1, from [23], with information on particle masses, spin, charges, and relationships with the bosons. (a fourth generation has effectively been excluded by experimental results [17,18,24,25]). Moving from generation 1 to generation 3 is achieved by increasing the mass of the previous generation

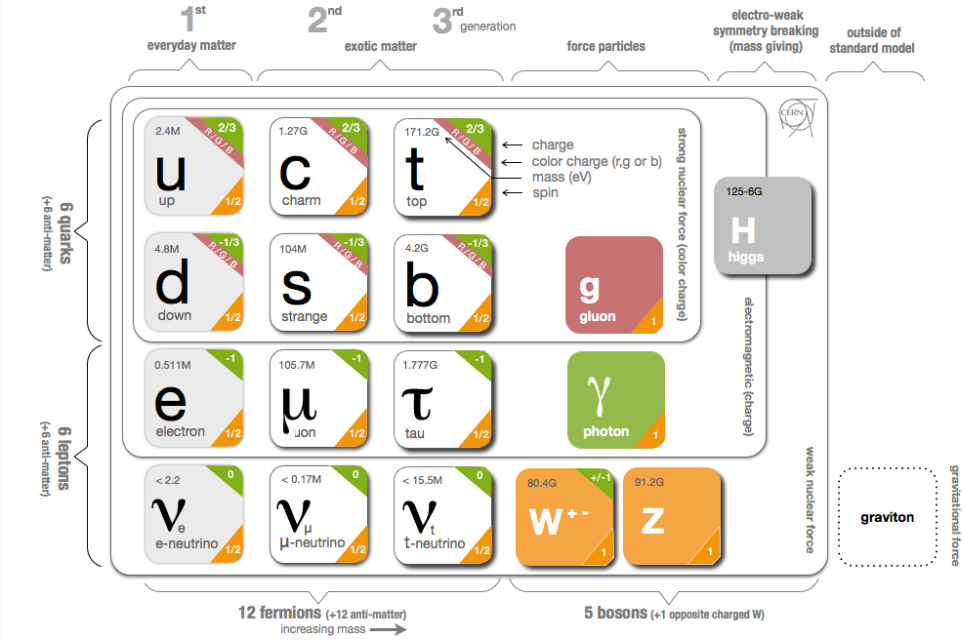

Figure 2.1: The elementary particles of the SM also shown with the hypothetical graviton.

to produce a new particle with identical quantum numbers. The generations are further split into two groups based on the interactions in which they can participate: quarks (all interactions) and leptons (singlets in $SU(3)$). Each particle has its own antiparticle with identical mass but opposite quantum numbers. Depending on the representation in $SU(2)$, fermions can exist as doublets (left-handed) or singlets (right-handed). Eq. 2.3 gives a general definition for the SM doublets:

$$\begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \quad \begin{pmatrix} q_u \\ q'_d \end{pmatrix}_L \tag{2.3}$$

where the lepton doublet contains neutrinos (ν_ℓ) and charged leptons ($\ell = e, \mu, \tau$) and the quark doublet includes up-type quarks ($q_u = u$,c,t) and down-type quarks ($q'_d = d'$,s',b'), where the prime indicates the weak eigenstates. The right-handed singlets are represented as $(\ell)_R$, $(q_u)_R$, and $(q_d)_R$ (no right-handed neutrinos have been experimentally observed).

Under $SU(2)$ transformations, the helicity of a given particle will affect its interactions. Unlike the Z boson and photon, the $W^\pm$ bosons only interact with $SU(2)$ doublets, Eq. 2.3. As such, interactions between $W^\pm$ and fermions result in changes in flavor, e.g., $\nu_{e,L} \to e_L$. While leptons only change flavor within the same generation, quarks are capable of changing flavors across generations. The Cabibbo-Kobayashi-Maskawa (CKM) matrix relates the

weak to the mass eigenstates, providing a mechanism for quarks to change flavor across generations, as shown in Eq. 2.4.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.4}$$

The value of each CKM element reflects the coupling of the up-type quark to the down-type quark, e.g., V_{tb} is the top-bottom quark coupling. The elements cannot be calculated in the SM, and instead must be measured. The non-diagonal elements have been measured to be small but non-zero [26].

2.1.1 Electroweak Symmetry Breaking

Each boson described in the previous section is generated without mass to preserve local gauge invariance. This consequence disagrees with many experimental observations that conclude there are massive bosons responsible for mediating the weak interaction [27–30]. The Brout-Englert-Higgs (BEH) mechanism resolved this discrepancy by generating mass for the electroweak bosons $W^{\pm}/Z$ [31–36]. For a complex scalar field represented as an SU(2) doublet,

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.5}$$

the potential energy term of the Lagrangian becomes $V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2$. The shape of the potential depends upon μ, the mass term, and λ, the four-way interaction vertex coupling. Choosing $\mu^2 < 0$ yields a non-zero result for the potential minimum at $\phi^\dagger \phi = -\mu^2/(2\lambda) \equiv v^2/2$. By performing an expansion around the minimum, the properties of the field ϕ can be explored. At the potential minimum, ϕ can be written as

$$\phi \to \phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.6}$$

such that only one component can be defined to be non-zero in ϕ ($\phi_3 = v$ in this definition). Expanding around the minimum with some real field $H(x)$ gives

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.7}$$

Propagating this new representation through the original Lagrangian, imposing local gauge invariance, correctly produces the massive gauges bosons $W^{\pm}/Z$ while maintaining a massless photon. Assigning one component of ϕ to be non-zero at the minimum spontaneously breaks the symmetry. In this framework, the minimum refers to the vacuum expectation value (vev), v. The vev is related to the mass of the particle in the field used to expand around the minimum $H(x)$, known as the the Higgs boson, by $m_H = \sqrt{2\lambda}v$, where v can be calculated from the measured $W^{\pm}$ mass. Yukawa couplings between the Higgs boson and fermions generates non-zero masses that can be measured. The coupling strength between the Higgs boson with fermions is proportional to the fermion mass, indicating that the heavier fermions, e.g., top and bottom quarks, couple most strongly with the Higgs. In 2012, the Higgs boson was discovered at the LHC experiments ATLAS and CMS [17,18] with a mass near 125 GeV, successfully completing the SM of particle physics.

2.2 Physics Beyond the Standard Model

Many experimental results, both collider and non-collider, have confirmed theoretical predictions across many orders of magnitudes and with high-precision. Despite the continued experimental validation of the SM, there are well-known discrepancies that have yet to be resolved. As a result, the SM is currently considered an effective field theory, requiring a theoretical extension to be complete. Much of the experimental effort at the Large Hadron Collider is directed at finding new physics that can provide explanations for the discrepancies, a few of which are described below.

Currently, as described in Eq. 2.1, the SM only describes the strong and electroweak interactions without any quantum description of gravity. Attempts to include the gravitational interaction in the SM have thus far been unsuccessful, limiting the scope of the SM to only three out of four known interactions.

Astrophysical observations indicate that a new kind of matter, approximately five times more abundant than the amount of matter predicted by the SM, exists in the universe. This new matter is referred to as *Dark Matter* (DM) as it is non-luminous and appears to only interact with SM matter through the gravitational force. No particle in the SM can be responsible for DM. Similarly, the observed accelerated expansion of the universe, *dark energy*, cannot be explained by the SM. In total, only $\sim$5% of the energy content of the universe can be explained by the SM.

Each particle of the SM has its own anti-particle and it is expected that the amount of matter and antimatter were produced in equal proportions. At present, the amount of antimatter is significantly less than the amount of matter visible in the universe. This

matter-antimatter asymmetry can only partially be explained by CP violation, requiring an extension to the SM to adequately parameterize this effect.

Finally, the quantum-loop corrections to the Higgs mass are predicted to quadratically diverge. However, the measured mass of the Higgs boson indicates that it is closer to the electroweak scale than the Planck scale. To prevent the divergence, either SM parameters are *fine-tuned* or new physics exists that *naturally* cancels the divergence.

Different theoretical extensions to the SM have been proposed to address many of these limitations. This dissertation explores a very simple extension to the SM, Vector-like Quarks (VLQs). VLQs are an appealing signature at the Large Hadron Collider because of their inclusion across many BSM theories and they offer a method to postpone the Higgs mass divergence.

2.3 Vector-like Quark Phenomenology

VLQs are a hypothetical class of spin–1/2 fermions that resemble the SM quarks in their $SU(3)$ representation with subtle differences that distinguish them from the hypothetical fourth generation fermions. Unlike the SM quarks, VLQs are non-chiral and thus predicted to have symmetric left- and right-handed couplings in the weak isospin $SU(2)$ group. The non-chiral VLQs do not require the BEH mechanism to generate mass, as SM quarks require Sect. 2.1. Furthermore, VLQs are currently the simplest colored fermion allowed by experimental data and, aside from the lack of a discovery, there is no theoretical reason why they shouldn't exist. SM measurements do limit the range of couplings between VLQs and SM, e.g., $Z \to b\bar{b}$ and others [37], but they do not yet exclude VLQs.

BSM scenarios, such as Little Higgs [38,39] and Composite Higgs [40,41] models, include VLQs in loop diagrams for Higgs production (gluon-gluon fusion) and Higgs decay ($H \to \gamma\gamma$) to (partially) cancel the quadratic divergence generated by radiative corrections to the Higgs mass. Each particular BSM framework derives a specific coupling between the VLQs and Higgs that affects the aforementioned processes.

The VLQ representation can include singlets, doublets, or triplets [42]:

$$\begin{array}{lll} T_{L,R}, \quad B_{L,R} & & \text{(singlets)}, \\ \Big(XT\Big)_{L,R}, & \Big(TB\Big)_{L,R}, \quad \Big(BY\Big)_{L,R} & \text{(doublets)}, \\ \Big(XTB\Big)_{L,R}, & \Big(TBY\Big)_{L,R} & \text{(triplets)} \end{array} \tag{2.8}$$

where T and B are similar to the top and bottom SM quarks (electric charge $+2/3$ and $-1/3$, respectively) and X and Y have more exotic electric charges, $+5/3$ and $-4/3$, respectively,

but similar properties as B and T. The results in this dissertation do not include any analysis of X and Y and will not be discussed further. VLQs are commonly referred to as 'top partners' because the large quark mass in the third generation, relative to the other generations, suggests VLQs will couple more strongly to the top and bottom quarks. The analyses presented here assume that the VLQs only couple to the third generation.

Extending the SM with VLQs introduces mixing between SM quarks and VLQs. In Eq. 2.9, the mixing between the top and VLQ T weak and mass eigenstates is described [43]. The same structure exists for the bottom quark and VLQ B, with couplings defined for down-type quarks.

$$\begin{pmatrix} t_{L,R} \\ T_{L,R} \end{pmatrix} = \begin{pmatrix} \cos\theta^u_{L,R} - \sin\theta^u_{L,R} e^{i\phi_u} \\ \sin\theta^u_{L,R} e^{-i\phi_u} \cos\theta^u_{L,R} \end{pmatrix} \begin{pmatrix} t'_{L,R} \\ T'_{L,R} \end{pmatrix} \tag{2.9}$$

A prime on the quarks indicates the weak eigenstate, θ is the mixing angle and ϕ the phase factor. For particular BSM interpretations, the exact value of θ depends on couplings between the SM quarks and VLQs.

$$\begin{aligned} \tan 2\theta^q_L =& \frac{\sqrt{2}|y^q_{34}|vm_0}{m_0^2 - |y^q_{33}|^2v^2/2 - |y^q_{34}|^2v^2/2} \quad \text{(singlets,triplets)}, \\ \tan 2\theta^q_R =& \frac{\sqrt{2}|y^q_{43}|vm_0}{m_0^2 - |y^q_{33}|^2v^2/2 - |y^q_{43}|^2v^2/2} \quad \text{(doublets)}, \end{aligned} \tag{2.10}$$

Eq. 2.10 represents the relationship between mixing angles and Yukawa couplings y^q_{ij}, $q = u, d$, that are defined in a particular BSM interpretation. The bare mass term for the VLQs, m_0, also alters the coupling definitions, along with the vev, v. The couplings between SM quarks and the SM electroweak bosons $W^\pm/Z/H$ are also affected by the mixing because the addition of VLQs modifies the elements of the CKM matrix. VLQ couplings to the SM electroweak bosons describes effects that are necessary for higher-order corrections, including the loop corrections to the Higgs boson mass.

Of critical importance to the analyses presented here is the mixing between SM quarks and VLQs, which allow the VLQs to decay. The Lagrangian can be written in general for the different couplings to quarks, VLQs, and boson, see Eq. 2.11.

$$\begin{aligned} \mathcal{L}_W =& -\frac{g}{\sqrt{2}}\bar{Q}\gamma^\mu(V^L_{Qq}P_L + V^R_{Qq}P_R)qW^+_\mu + \text{h.c.} \\ & -\frac{g}{\sqrt{2}}\bar{q}\gamma^\mu(V^L_{qQ}P_L + V^R_{qQ}P_R)QW^+_\mu + \text{h.c.} \\ \mathcal{L}_Z =& -\frac{g}{\sqrt{2c_W}}\bar{q}\gamma^\mu(\pm X^L_{qQ}P_L \pm X^R_{qQ}P_R)QZ_\mu + \text{h.c.}, \\ \mathcal{L}_H =& -\frac{gm_Q}{2m_W}\bar{q}(Y^L_{qQ}P_L - Y^R_{qQ}P_R)QH + \text{h.c.} \end{aligned} \tag{2.11}$$

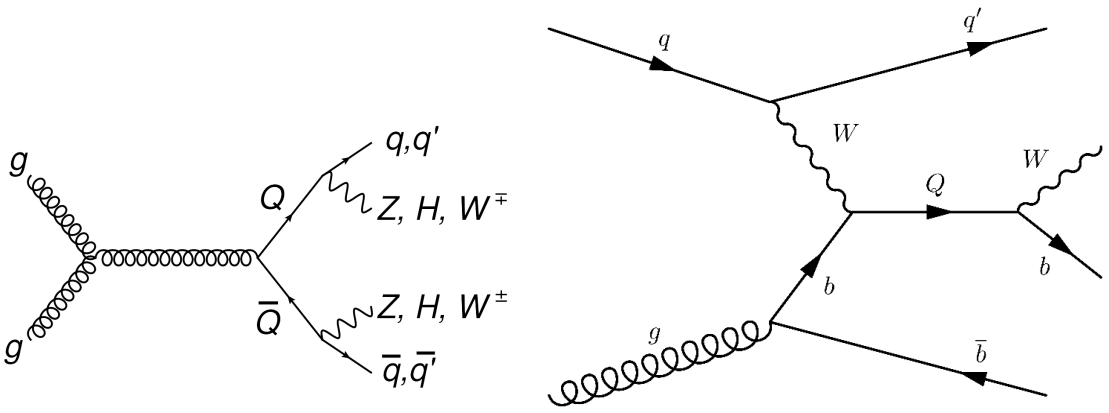

Figure 2.2: Feynman diagrams for the pair production (left) and single production (right) of VLQs at the LHC.

where $P_{R,L} = (1 \pm \gamma_5)/2$, g is the weak coupling constant, and c_W is the weak mixing angle. The specific couplings ($V^{L,R}$, $X^{L,R}$, $Y^{L,R}$) are given by the elements of the (extended) CKM matrix [43]. BSM theories are differentiated by the predicted branching ratios for VLQ decays. A discovery will require a measurement of the VLQ branching ratios to understand the couplings in Eq. 2.11 and to discriminate between BSM theories.

At the Large Hadron Collider, VLQs can be produced individually or in pairs. Pair production is considered more model independent than single production. The gauge coupling between VLQs and gluons only depends on the VLQ mass for pair production while the single production depends on the electroweak parameters at the boson-quark-VLQ vertex, diagrammed in Fig. 2.2. For large enough masses of VLQs, the single production mechanism can become more favored than pair production at the LHC, as highlighted in Fig. 2.3 where the cross section values are calculated using `top++` [44]. The point at which single VLQ production dominates is model-dependent, but larger masses will be more favorably produced with the center of mass energy currently accessible.

From Eq. 2.11, the coupling of VLQs to SM particles occurs through Yukawa couplings. As such, VLQs are capable of decaying into a boson and quark, e.g. shown in Fig. 2.2 and described in detail in [45]. The branching ratios for $SU(2)$ singlet and $SU(2)$ doublet scenarios, as a function of VLQ mass as computed with PROTOS [46,47], are shown in Fig. 2.4. The resulting VLQ phenomenology at the LHC is quite rich and offers many opportunities to explore high energy physics object reconstruction and identification in a unique topology.

This dissertation explores two different, but complementary, searches for VLQs at 13 TeV with the ATLAS detector. The first search explores the single lepton final state, focusing on $T\bar{T} \to Wb + X$, where $X = Wb, Zt, Ht$ [48]. The second search is more general and uses the fully-hadronic final state (no leptons) to search for all decay modes of $T\bar{T}/B\bar{B} \to$ jets.

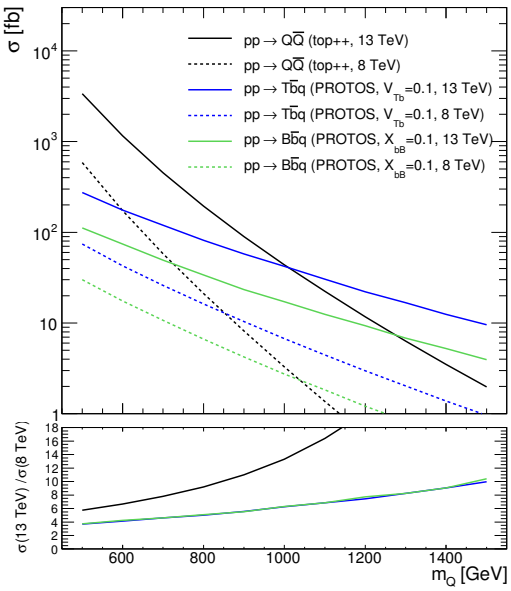

Figure 2.3: Relative production cross sections at $\sqrt{s} = 8/13$ TeV as a function of VLQ mass for single and pair production. Cross section values are calculated using `top++`.

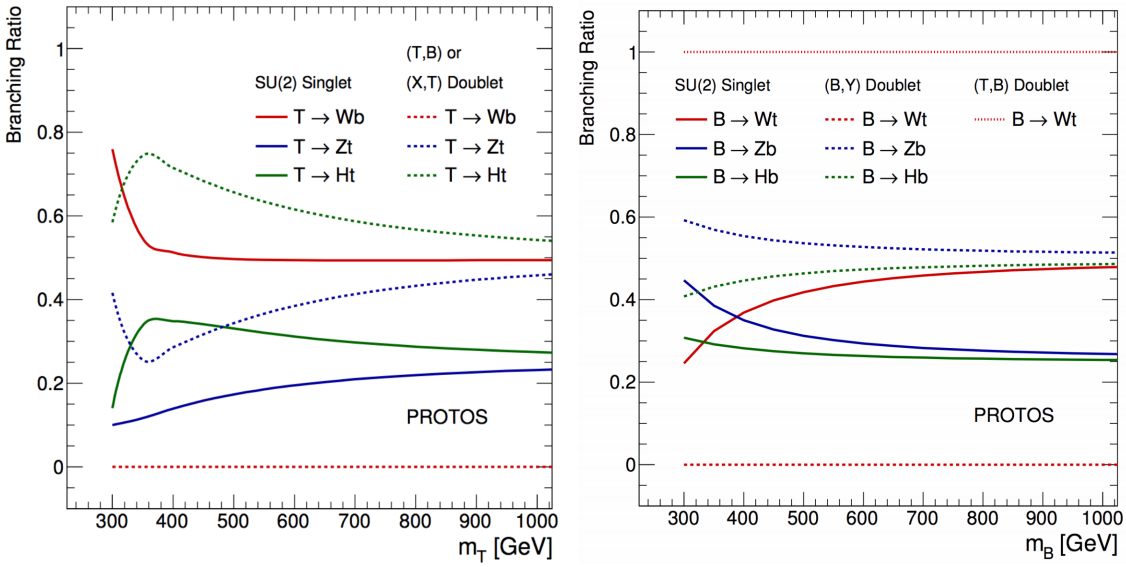

Figure 2.4: Branching ratios for T (left) and B (right) as a function of VLQ mass.

CHAPTER III

The ATLAS Experiment at the Large Hadron Collider

3.1 The Large Hadron Collider

The CERN accelerator complex contains the world's largest and highest energy particle accelerator, the Large Hadron Collider (LHC) [49]. Two beams of protons are accelerated in opposite directions and collide at a center-of-mass energy of 13 TeV.

A multi-stage accelerator system, shown in Fig. 3.1, is responsible for acquiring protons from hydrogen atoms and ultimately colliding them at a center-of-mass energy of 13 TeV. Protons are initially accelerated to 50 MeV in the LINAC2. Next, the protons transfer to the Proton Synchrotron Booster, a circular accelerator 50m in diameter, where they accelerate to 1.4 GeV. The Proton Synchrotron Booster injects the protons to the Proton Synchrotron, a circular accelerator with a 200m diameter, where an energy of 25 GeV is achieved. From the Proton Synchrotron the protons continue into the Super Proton Synchrotron, another circular accelerator with a diameter of nearly 2300m that accelerates the protons to 450 GeV. The Super Proton Synchrotron is the final stage before protons enter the LHC. Two separate beams of protons enter the LHC heading in opposite directions and continue to accelerate along the 27 km ring until they reach their peak energy of 6.5 TeV and collide at the interaction points. The proton acceleration in the LHC is achieved using two independent, superconducting Radio Frequency (RF) systems, one for each proton beam, operating at 400 MHz. Each RF system consists of eight cavities, powered by a single klystron to improve control of the field in each cavity, yielding a higher-quality beam structure for reliable proton–proton collisions. Because the LHC exists in the same tunnel that contained the Large Electron-Positron collider, it was not possible to construct two separate beam pipes for each circulating proton beam. As such, a twin-bore magnet accelerates two beams of protons in opposite directions within the same cryostat. Superconducting dipole and quadrupole magnets are responsible for steering and focusing the proton beams, respectively.

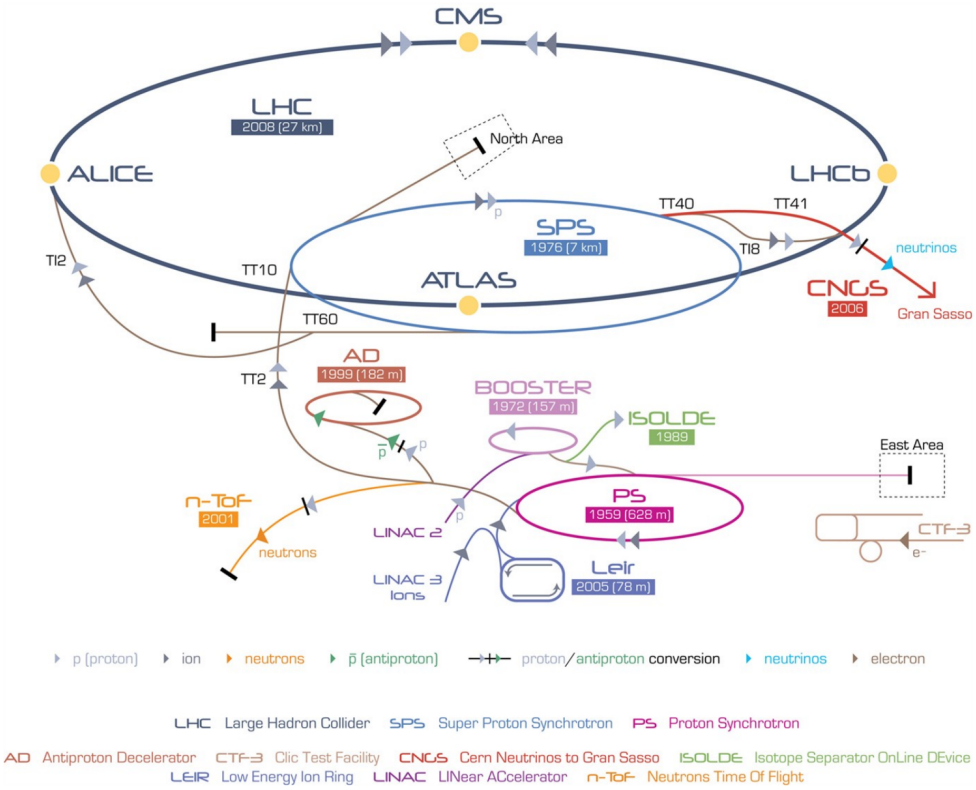

Figure 3.1: Schematic of the LHC accelerator complex. The particle detectors and stages of the proton acceleration are labeled.

Protons are circulated in distinct groups, referred to as bunches, and focused by the quadrupole magnets at the interaction points. By focusing the bunches, the cross-sectional area of overlap, $\mathcal{A}$, of each bunch is reduced and increases the instantaneous luminosity, $\mathcal{L}$, defined in Eq. 3.1. The luminosity is a key performance metric of the LHC for its associated experiments. A large luminosity increases directly increases the production of SM and BSM physics processes, thus improving the sensitivity for the analyses at the LHC experiments. The instantaneous luminosity also depends on the number bunches, n_b, the number of protons in each bunch, N, and the frequency, f, at which the bunches are collided.

$$\mathcal{L} = f n_b \frac{N_1 N_2}{\mathcal{A}} \tag{3.1}$$

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN_{\text{tot}}}{dt} \tag{3.2}$$

The cross-sectional area of overlap can be described using the two properties emittance and β^*. The emittance is a representation of the distribution of protons within the bunch and β^* describes the amplitude function, i.e., beam optics at the interaction point. The emittance and beta function are currently limited by the capabilities of the existing magnet system. The revolution frequency and number of protons per bunch can be increased to proportionally increase the luminosity. However, there are accelerator and detector design capabilities that limit the revolution frequency and number of protons per bunch. The luminosity provided by the LHC for 2015 and 2016 is summarized in Figs. 3.2 and 3.3. In 2015, the LHC achieved a peak luminosity of 5×10^{33} cm^{-2}s^{-1} and total integrated luminosity of 4.2 fb^{-1}, of which 3.9 fb^{-1} was successfully recorded by ATLAS. The performance of the LHC was significantly improved for 2016, reaching a peak luminosity of 1.38×10^{34} cm^{-2}s^{-1}, greater than the designed luminosity of 1×10^{34}, and a total integrated luminosity of 38.5 fb^{-1}, with 35.6 fb^{-1} recorded by ATLAS. Table 3.1 lists the current operating parameters of the LHC. For each of these parameters, the LHC is currently operating near design values.

3.2 The ATLAS Experiment

The ATLAS detector is one of two multi–purpose particle detectors designed for sensitivity to discovering new physics and performing precision measurements of SM properties [50]. ATLAS is a cylindrical detector with a diameter of 25 m and a length of 44 m. At a center-of-mass energy of 13 TeV, the ATLAS experiment can probe distances at the scale of 10^{-20}m. ATLAS uses information collected by four detector subsystems to achieve this level of performance. The subsystems are graphically represented in Fig. 3.4 with example interactions

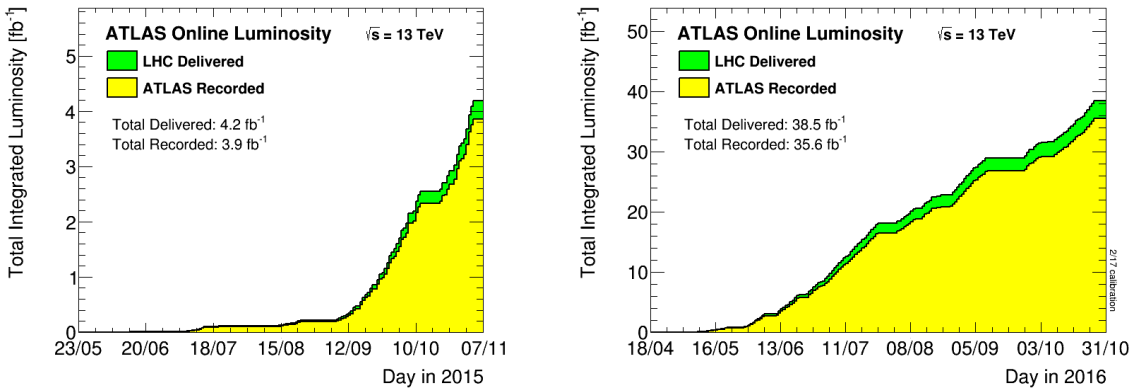

Figure 3.2: Integrated luminosity recorded by the ATLAS detector for each day in 2015 (left) and 2016 (right).

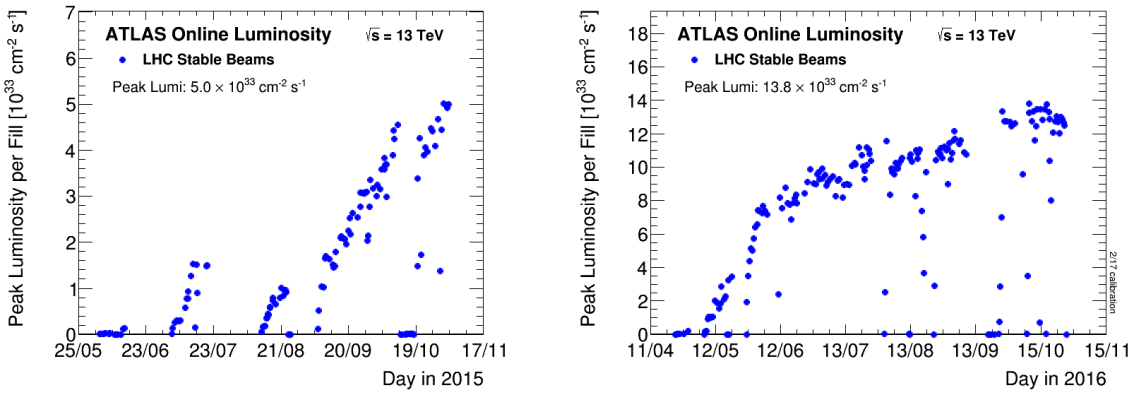

Figure 3.3: Peak luminosity recorded by the ATLAS detector for each day in 2015 (left) and 2016 (right).

Table 3.1: LHC beam parameters for 2015 and 2016. Values retrieved from the LHC Performance Workshop (Chamonix 2017). The parameter $\langle\mu\rangle$ is the mean number of interactions per bunch crossing.

Parameter	2015	2016
beam energy [TeV]	6.5	6.5
β^* [m]	0.8	0.4
bunch spacing [ns]	25	25
number of bunches	2244	2220
protons per bunch (10^{11})	1.2	1.2
peak luminosity [10^{34} cm^{-2}s^{-1}]	0.5	1.38
$\langle\mu\rangle$	13.7	24.9

from SM particles and the full detector is depicted in Fig. 3.5.[1] From closest to farthest away from the interaction point, the detector subsystems are: an inner detector to identify charge particles (Sect. 3.2.1), a powerful 2 T solenoid magnet to distinguish between charged particles, an electromagnetic calorimeter to measure energy deposited by electrons and photons and a hadronic calorimeter that measures energy deposited by hadrons (Sect. 3.2.2), and finally a muon spectrometer (Sect. 3.2.3) surrounded by three superconducting toroid magnets. The trigger system (Sect. 3.2.4) is responsible for selecting which events are stored for analysis.

3.2.1 Inner Detector

The inner detector (ID) is responsible for identifying charged particles in the central region of the detector $|\eta| < 2.5$. The ID has high momentum and spatial resolution such that individual vertices can be reconstructed for each collision.

The ID is comprised of four separate subsystems, the insertable B-layer (IBL), pixel layer, silicon microstrip tracker (SCT), and the transition radiation tracker (TRT). Each of these layers and their relative position to the beam pipe are shown in Fig. 3.6. The entire ID resides inside a 2 T magnetic field that adds charge identification information to reconstructed physiscs objects. The subsystems are designed and built with considerations to withstand the large amounts of expected radiation (radiation hardness), hardware maintenence, and physics performance.

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. [51]

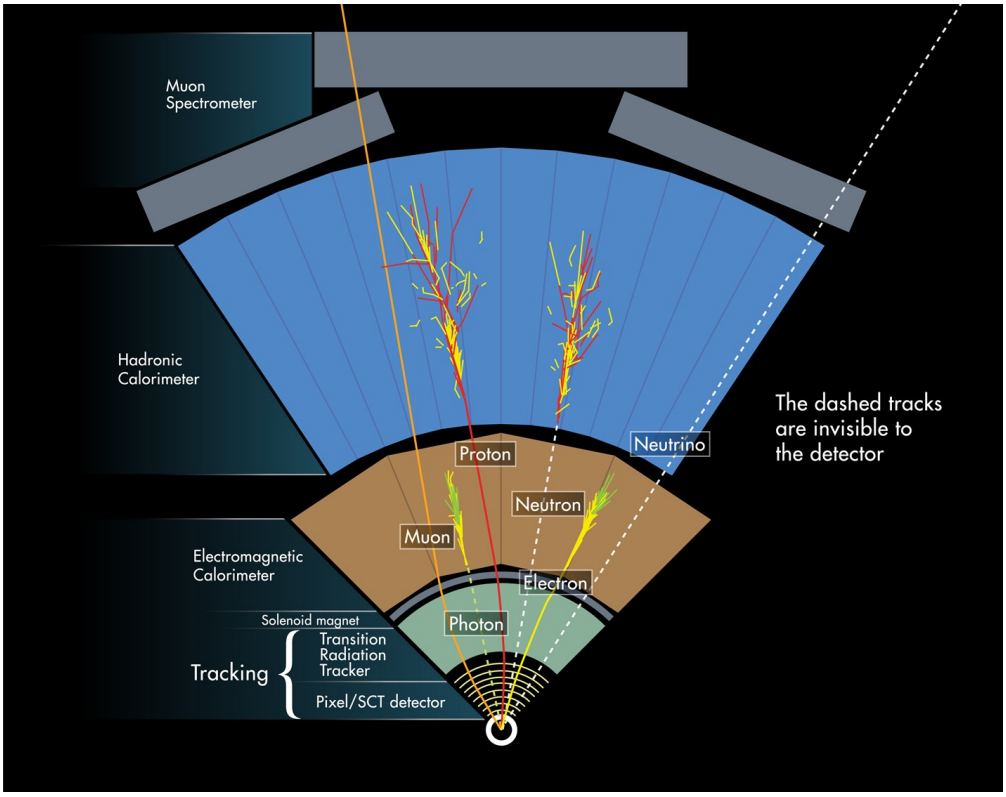

Figure 3.4: Graphical representation of the ATLAS detector technology. Characteristic interactions of different particle species with the various subsystems are shown.

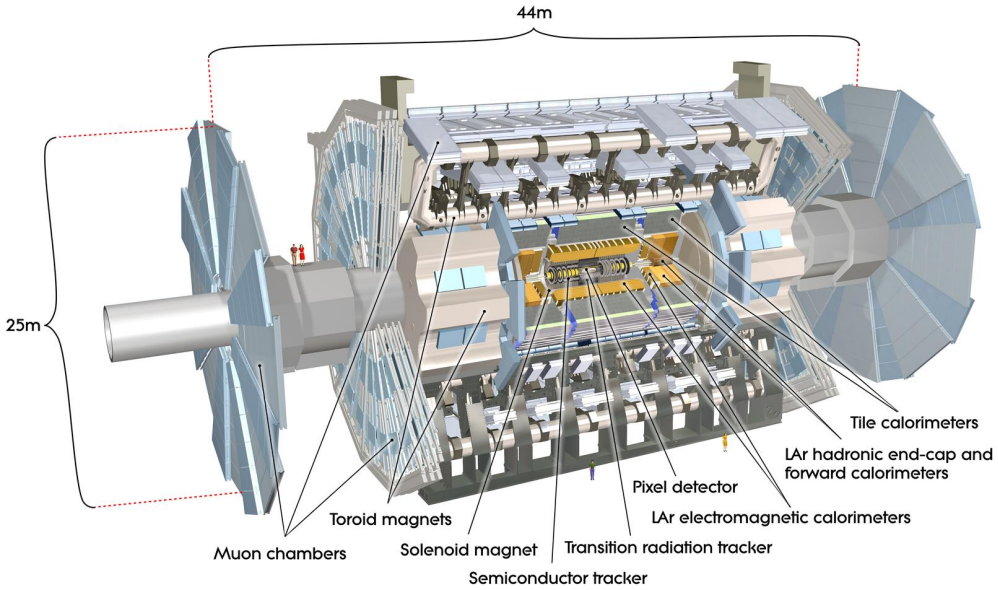

Figure 3.5: Graphical representation of the ATLAS detector. Each subsystem is labeled and the relative scale is compared to humans shown at the bottom and left-side of the detector.

The IBL is the closest detector component to the interaction point, vastly improving vertex reconstruction and identification over the run 1 performance [52]. Added between runs 1 and 2, the IBL is a cylindrical detector inserted between the beam pipe, reduced in diameter after run 1 specifically for the IBL, and the pixel layer. The IBL detector preserves reconstruction performance for high-luminosity conditions provided by the LHC in run 2 and future runs that would compromise the integrity of the pixel layer performance. The IBL pixels have an area of 250×50 μm^2 and a space-point resolution per module $< 10\mu$m in $r - \phi$ and 72μm in z.

The pixel layer consists of 3 layers of silicon pixels at $r = 50.5$, 88.5, 122.5 mm from the interaction point in the barrel region. The endcap region, where the detector is oriented radially and perpendicular to the beam line, also has 3 pixel layers at $r = 495$, 580, and 650 mm. All of the pixel layer sensors are identical and the space-point resolution of 14μm^2 in $r - \phi$ and 115μm in z (r) for the barrel (endcap) offers excellent resolution for charge particle reconstruction and identification. The SCT consists of four layers of silicon microstrip detectors in the barrel region and nine layers in the endcap region. The barrel region SCT layers range from $r = 299$ mm to $r = 514$ mm and the endcap layers begin at $r = 852.8$ mm and end at $r = 2720.2$ mm. The TRT contains layers of polyimide straw tubes, each 4 mm in diameter and 144 cm long, filled with a gaseous mixture of xenon (70%), CO_2 (27%), and CF_4 (3%). There are 73 layers in the barrel region extending from 554 to 1082 mm in r, and 160 layers in the endcap, covering 848 to 2710 mm in r. Unlike the other ID subsystems, the TRT relies on charged particles ionizing the gas inside the straw tubes to generate electrical signals. The ionized electrons inside the gas take up to 48 ns to reach the anode wire in the center of the tube and yields a position resolution of 130μm.

3.2.2 Calorimeters

The calorimeter system is located between the ID and muon spectrometer. This system is designed to detect all particles that undergo electromagnetic interactions (electrons or photons) or hadronic interactions (hadrons).

Particles that interact with material in the calorimeters will generate cascades of secondary particles referred to as *showers*. Depending on the interaction, the characteristics of the showers within the calorimeter can vary. The calorimeter system consists of an electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). Both calorimeters are sampling in nature, consisting of alternating layers of a dense absorber and active material. The dense absorber layer serves to stop particles and initiate showers and the active material measures the energy produced in the shower. To accurately reconstruct the energy of the initial particle, the sampling fraction of the material must be known to rescale the

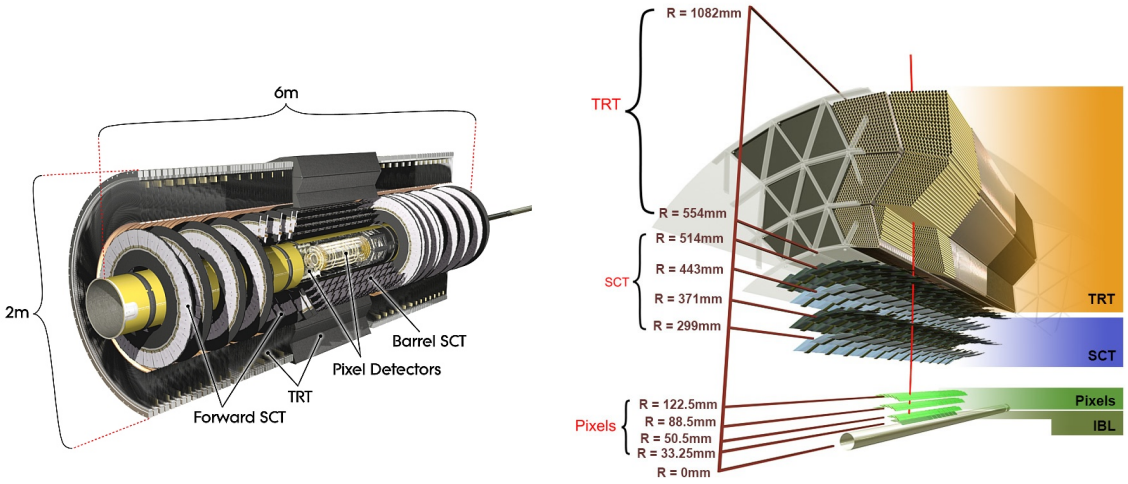

Figure 3.6: The ATLAS inner detector full schematic and labels (left) and barrel region with radial positions (right).

measured energy due to energy lost in the absorber. The relative positioning and size of the calorimeters are shown in more detail in Fig. 3.7.

Due to its radiation hardness, uniform structure, and response stability, liquid argon (LAr) was chosen as the ECAL active material. Lead serves as the dense absorber material in the ECAL because of its short radiation length $X_0 = 0.561$ cm that improves containment of the showers within the ECAL volume. Copper electrodes are arranged in an alternating accordion geometry with the lead absorbers, diagrammed in Fig. 3.7, and immersed in the LAr to ensure full coverage in ϕ. There are three regions of varying granularity in the ECAL to enable precise reconstruction in the initial layers and simply capture the remainder of the electromagnetic shower in the final layer. Particle showers ionize the LAr and a high-voltage applied to the copper electrodes forces the electrons to drift onto the plates. The electrical signal is too long to be completely measured by the detector given the 25 ns revolution frequency of the LHC, the full signal arrives after approximately 450 ns, but a majority of the information arrives quickly and decays linearly to zero. To compensate, the signal from the ECAL is shaped into a bipolar pulse signal with a much shorter readout time designed to match the 25 ns bunch spacing. and then used to measure the deposited energy. Both the barrel (EMB), $|\eta| < 1.475$, and endcap (EMEC), $1.37 < |\eta| < 3.2$, incorporate this detector technology, with a separate cryostat for each region. The transition region between the EMB and EMEC, $1.37 < |\eta| < 1.52$, contains extra material, resulting in degraded physics performance and is not used in the results presented here. A LAr presampler is placed before the ECAL to correct for energy lost prior to the calorimeter, extending up to $|\eta| < 1.7$.

The detector technology implemented for the HCAL varies between barrel and endcap. The barrel region is covered by two separate HCAL regions, the barrel and extended barrel,

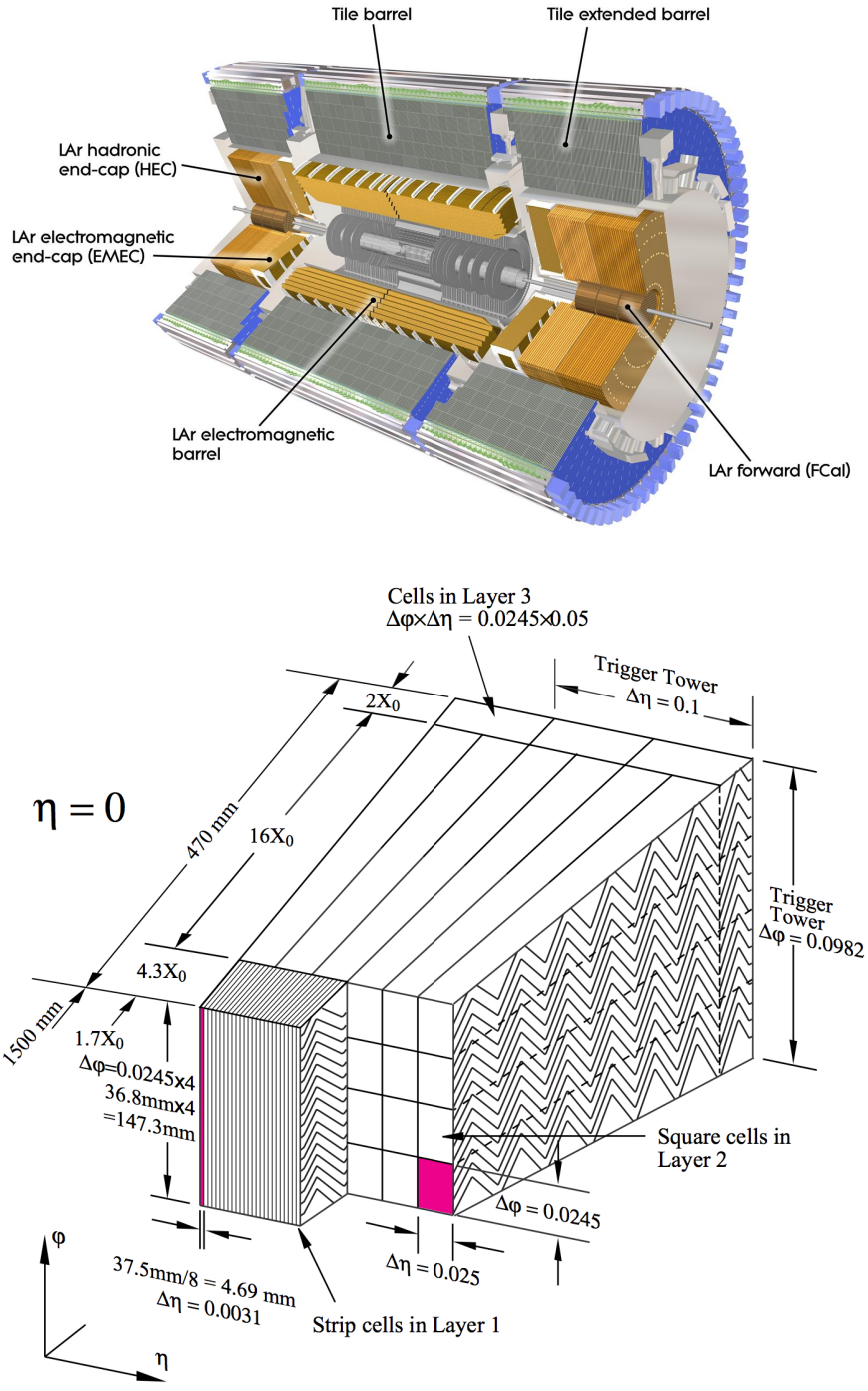

Figure 3.7: Full calorimeter system with different regions labeled (top) and schematic of the electromagnetic calorimeter segments with dimensions and regions labeled (bottom).

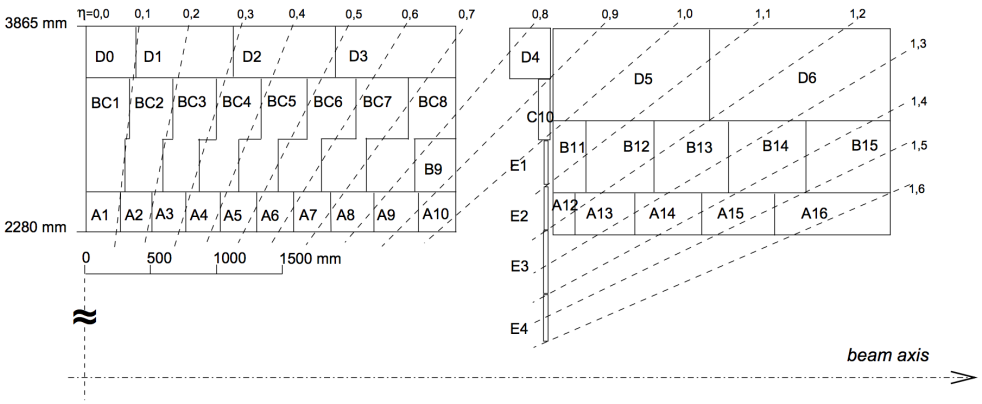

Figure 3.8: Schematic of the hadronic endcap and barrel calorimeters with positioning of different segments.

that use steel absorber and scintillating plastic active regions. The barrel HCAL covers $|\eta| < 0.9$ and the extended barrel HCAL covers $0.8 < |\eta| < 1.7$. Photomultiplier tubes detect light produced in the scintillating tiles from hadronic interactions. The hadronic endcap calorimeter (HEC) uses LAr as the active medium, similar to the EMEC. The HEC uses copper instead of a lead absorbing material and the accordion shape is abandoned for this detector. The range $1.5 < |\eta| < 3.2$ is covered by the HEC, and it is placed in the same cryostat as the EMEC. The full range and structure of the HCAL barrel detectors are shown in Fig. 3.8.

To cover the region $|\eta| > 3.2$, a special set of forward calorimeters (FCAL) is required that can withstand high-levels of radiation. The FCAL detector extends from $3.1 < |\eta| < 4.9$ and is broken into 3 regions, each using LAr as the active medium. The electromagnetic part of the FCAL uses copper as the absorber, similar to the HEC, and the hadronic portions of the FCAL use tungsten. Drift times are significantly reduced in the FCAL compared to the ECAL by reducing the separation between electrodes.

3.2.3 Muon Spectrometer

At the LHC, muons are minimum-ionizing particles and pass through the ID, ECAL, and HCAL with minimum energy deposition. The last detector subsystem, the muon spectrometer (MS) includes two separate detectors, in the barrel and endcaps up to $|\eta| < 2.7$, designed specifically to detect muons, depicted in Fig. 3.10. Complementary technologies are chosen through the MS to provide the best position and timing resolution for muon reconstruction. In the barrel region, monitored drift tubes (MDTs), which are similar to the TRT with larger tube diameters, have a position resolution of $\sim$80 μm and a maximum drift time of $\sim$700 ns. Resistive plate chambers (RPCs) are a complimentary technology in the barrel that operate

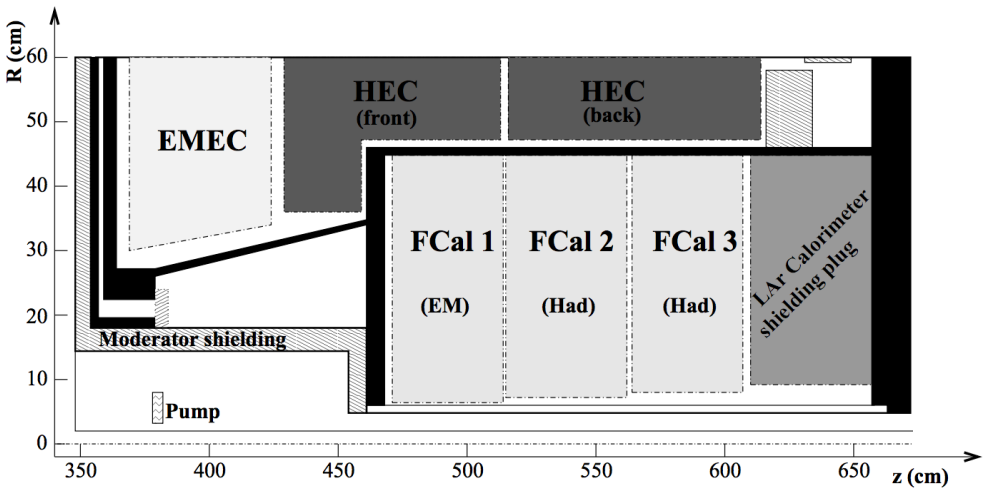

Figure 3.9: Forward calorimeter schematic depicting the different segments of the detector and positioning relative to the electromagnetic and hadronic endcap calorimeters.

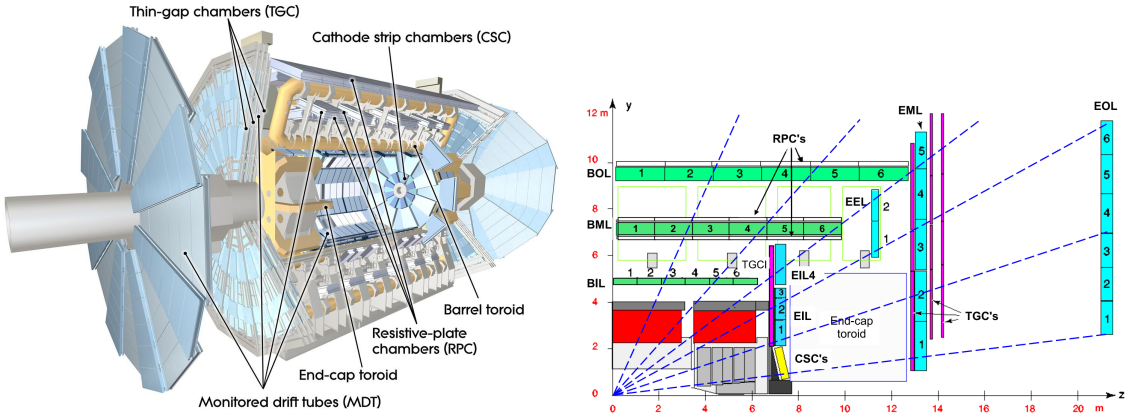

Figure 3.10: Muon spectrometer.

on shorter timescales of about 15-25 ns with worse position resolution of approximately 10 mm in z and r. Cathode strip chambers (CSCs) in the endcap only provide coverage from $2.0 < |\eta| < 2.7$, with a radial resolution of 40μm and time resolution of 40 ns. Beyond the CSCs, thin gap chambers (TGCs) operate in a similar η range and with similar characteristics as the RPCs. The CSC produces signals with 15-25 ns with spatial resolutions of 2-6 (3-7) mm in r (ϕ). The entire MS is immersed in a magnetic field produced by barrel and endcap toroids to improve muon identification. The toroidal fields have a bending power [2] of 1.5 to 5.5 Tm for $|\eta| < 1.4$ and 1 to 7.5 Tm for $1.6 < |\eta| < 2.7$.

[2]Bending power refers to the field integral $\int Bdl$, where B is the field component normal to the muon direction and the integral is computed along an infinite-momentum muon trajectory inside the muon chambers.

3.2.4 Trigger system

The ATLAS detector cannot record all of the information from the various sub-detectors for the collisions produced by the LHC every 25 ns (40 MHz). Additionally, not all collisions produce 'interesting' physics events, where the definition of interesting can vary between different physics processes. To overcome the physical limitations and ensure all interesting processes are saved, ATLAS implements a three-level trigger system with increasing levels of detector information to select events [53].

The Level 1 (L1) trigger is hardware-based and uses only coarse resolution information from the calorimeters and MS to reduce the rate to below 100 kHz. Information from the event is passed to the next level, Level 2 (L2) to be analyzed. Additionally, regions-of-interest are identified within the detector in η and ϕ for L2 to use in the trigger decision.

After passing the L1 trigger, the event information is passed to the final step in the trigger chain, a software-based trigger referred to as the high-level trigger (HLT). Either a full reconstruction of the event, similar to an offline analysis, or the full granularity of the detector within the regions-of-interest defined in L1 are used in the HLT trigger. The rate is ultimately reduced to the maximum readout of 1 kHz after the HLT. If an event passes the HLT, the full ATLAS data is recorded for analysis. In this dissertation, only HLT triggers are considered for selecting events. More information about the selected triggers can be found in Sects. 5.1.2 and 5.2.2

A wide variety of triggers are implemented by the ATLAS collaboration. Some triggers do not follow the aforementioned chain to acquire events to be studied in a performance context, e.g., calibration events. Triggers can be *pre-scaled* to accept a specific fraction of events for regions where the rate is too high for the trigger system. With pre-scaled triggers, analyzers can study regions of phase space inaccessible with the full trigger system, at the cost of a reduced luminosity. In this dissertation, pre-scaled triggers are not considered.

CHAPTER IV

Experimental Methods

4.1 Simulation of pp Collisions

In pp collisions, the final state is not known prior to the collision. Instead, a particular final state is produced according to its cross section $\sigma(pp) \to X$. To study a particular final state, it is possible to isolate data events that resemble the desired final state, but it cannot be known with absolute certainty what process the event actually represents. To ensure that events of a particular process can be studied, analyzers use Monte Carlo (MC) simulations of $pp \to X$ events. MC simulations give analyzers the ability study and understand the impact of the ATLAS detector on measured quantities, e.g., jet p_{T}, lepton charge, $E_{\mathrm{T}}^{\mathrm{miss}}$, etc. The specific generators used in the analysis are described in Sects. 5.1.1 and 5.2.1.

MC simulations use factorization theorems to consider different sub-processes of the pp collision, shown in Fig. 4.1, independently in simulation. A given MC simulation is produced using single or multiple generators to model a physical processes.

The scattering of two partons in the pp collision that results in a large transfer of momentum is known as the *hard scatter*. The hard scatter is first modeled in MC simulations and relies on the QCD production cross section for each process X, $pp \to X$, to relate the initial and final state particles in the collision. This step also considers the parton distribution functions (PDFs) to appropriately sample the momentum fractions and flavors of interacting partons. The specific cross section for a given process σ is calculated to a specific order in perturbation theory using the PDF information and collision energy. At this level, the partons in the final state, e.g., top quarks from a $gg \to t\bar{t}$ process, exist only as simulated particles and are referred to as *truth* particles.

After the simulation of the hard scatter, the MC simulation produces the *parton shower*. Partons in the initial and final state of the collision can radiate gluons, photons, and bosons, depending on the characteristics of the process. These extra radiations are considered as higher-order corrections to the hard scatter process, extending the $2 \to n$ process to $2 \to$

$n + X$. Each new parton added in the process is showered, e.g., $q \to qg$, according to a probability relating the invariant mass of the partons produced from the showering and the hadronization scale. The showering process continues until there is no more splitting of partons and then the hadronization step begins.

After showering, the final step of the MC simulation results in partons reaching an energy scale $\simeq$1 GeV and hadronizing to form color-less bound states. At this scale, perturbative QCD is no longer applicable and phenomenological models are required to accurately describe the hadronization process. These phenomenological models including the Lund string model and cluster model have parameters tuned to existing data to best describe the hadronization process [54].

As described in Sect. 3.1, protons circulate the LHC in bunches. Multiple protons can interact at the collision point and produce deposit energy in the detector. The soft interactions of partons from the same proton as the hard scatter are referred to as *underlying event*, and the partons from other protons are referred to as pile-up. Similar phenomenological models are used to simulate both effects. Pile-up can be produced *in-time* for protons that interact in the same bunch crossing and *out-of-time* for protons that interact in a different bunch crossing but is processed with the current bunch crossing due to electronic read-out.

After the MCs have generated the truth particles and initiated the parton shower, the four-vectors of all truth particles are passed to the detector simulation, alongside simulations of pile-up and underlying event, to produce reconstruction-level results. The ATLAS detector simulation includes simulated interactions of truth-level particles with detector materials using `GEANT4` [55, 56], converting deposited energy into electrical signals, and simulating the read-out electronics. The official detector geometry is documented for each pp collision run and any changes will result in modifications to the simulated detector geometry. For each modification to the detector geometry, simulated samples are re-processed to ensure an accurate representation of data-taking conditions is available to the analyzers. The information at reconstruction-level is processed identically to processed data, allowing the simulations to easily propagate through an analysis chain. Additionally, the truth-level information remains and can be used for different studies including calibrations, optimization, efficiencies, and signal acceptance.

4.2 Reconstructed Physics Objects

Physics and performance analyses with the ATLAS experiment require the translation of the electrical signals in the various detector subsystems into physics objects. The physics

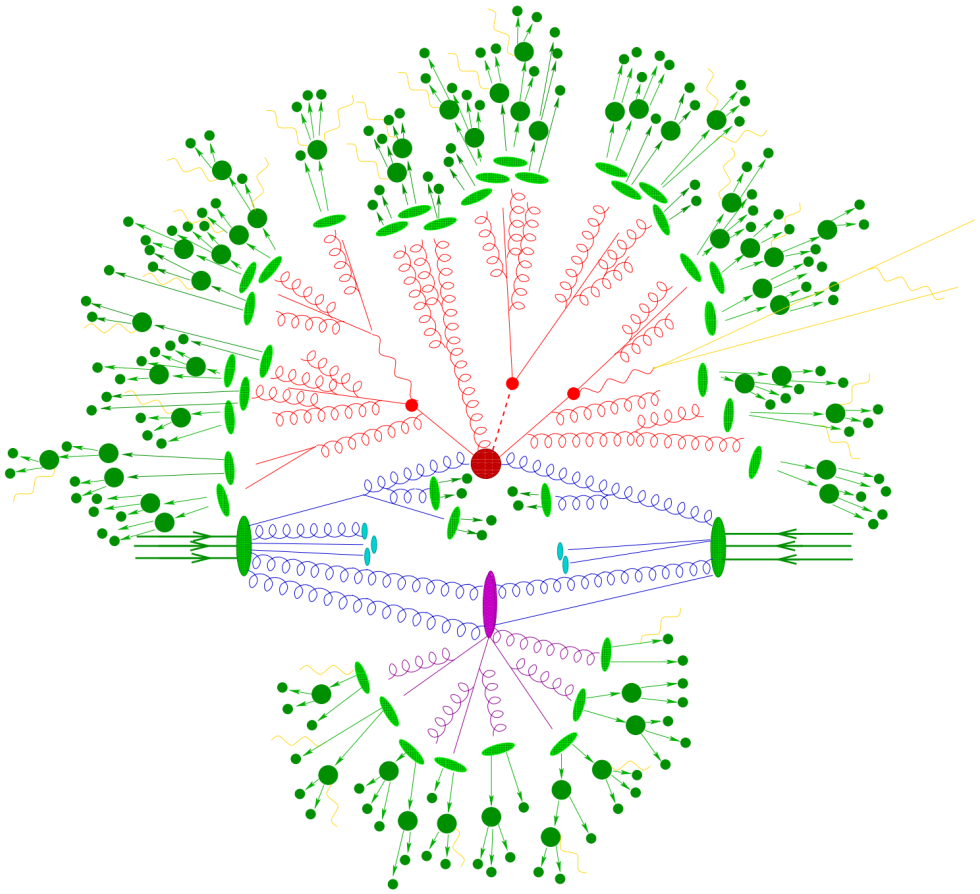

Figure 4.1: Schematic of a simulated $pp \to t\bar{t}H$ event [1]. The hard scatter is produces the top quarks and Higgs bosons (red circles) and additional QCD radiation (red). The final state partons produced from the $t\bar{t}H$ decays hadronize (light green) and decay (dark green). A secondary interaction occurs (purple) and photon radiation occurs at multiple stages (yellow).

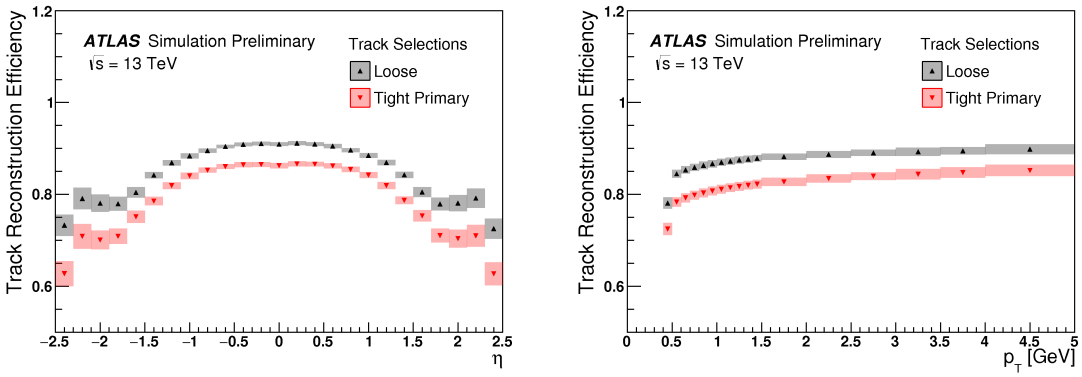

Figure 4.2: Track reconstruction efficiency for loose and tight identification as a function of η (left) and track p_{T} (right).

objects become powerful tools in an analysis to identify new physics or precisely measure SM properties.

4.2.1 Tracks

As charged particles move through the ID, energy is deposited within each layer, the IBL, pixel, SCT, and TRT. By combining the different hits within the layers of the ID using pattern recognition algorithms, the complete four-vector information of particle tracks is reconstructed [57]. Here, detector hits refers to measured signals produced in the detector during a pp collision. A fitting procedure tests the compatibility between the reconstructed tracks with predicted track patterns developed under certain hypotheses (pion or electron) to quantify the quality of the track reconstruction [58]. Additionally, the magnetic field generated by the solenoid bends particles in the ID, allowing for the discrimination between positively and negatively charged objects. Based on the number of hits in different layers of the ID, tracks are categorized into two groups, loose and tight, to reduce contamination from backgrounds. The η and p_{T} dependence on the efficiency for track reconstruction in these two categories is shown in Fig. 4.2.

Interaction points, known as primary vertices PVs, are identified using the reconstructed tracks. The PVs are crucial to the reconstruction of physics objects for matching the objects associated with the hard scatter. An iterative procedure matches tracks to PV candidates and discards tracks that are not significantly associated with that vertex. The discarded tracks are then used to fit a new primary vertex. This process continues until at least two tracks are matched to each PV. The PV defined to be the hard scatter has the largest sum of p_{T}^2 over all associated tracks.

4.2.2 Electrons

Inputs from both the ECAL, HCAL, and ID are used to reconstruct and identify electrons within the detector region $|\eta| < 2.47$ of the ATLAS detector [59]. The analyses presented here use electrons for triggers and $T\bar{T}$ reconstruction, see Sect. 5.1, or to veto events, see Sect. 5.2. Multiple steps are considered for electron reconstruction, beginning with seed-cluster reconstruction [60]. In the $\eta - \phi$ space, a sliding window compatible with the ECAL granularity, a size of 3×5 in units of 0.025×0.025, searches for deposits of energy within the window that contain $E_T > 2.5$ GeV. If duplicate seeds are identified, the seed with the smaller transverse energy is discarded. ID tracks, extrapolated to the middle layer of the ECAL, are matched in the $\eta - \phi$ space to the ECAL seed's barycenter, with effects from bremsstrahlung and precision hits in the SCT considered in the matching requirements. If no tracks are matched to a seed, the cluster is discarded from electron reconstruction. If more than one track is matched to a seed, a primary track is identified considering the distance to the seed barycenter, number of pixel and SCT detector hits. After matching at least one track to a calorimeter seed, the cluster is re-formed using a larger window of 3×7 (5×5) in the barrel (endcap). Once the reconstruction is complete, the energy associated to electron candidates is calibrated considering information from the presampler and an MC-based analysis. The electron position information, η and ϕ, is calculated using the primary track while the electron energy is determined by its ECAL cluster. To reduce the amount of non-prompt electrons, e.g., hadronic jets or photons, the electron's primary track must be associated with the primary vertex, satisfying $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma_{d_0}| < 5$.

Electron candidates are identified using a MC-based likelihood ratio $\mathcal{L}$ to discriminate between prompt (signal) and non-prompt (background) electrons:

$$\mathcal{L} = \frac{L_S}{L_S + L_B}, \quad L_{S,B}(\vec{x}) = \prod_{i=1}^{n} P_{(S,B),i}(x_i) \tag{4.1}$$

where $\vec{x}$ represents the vector of all 20 input variables used to discriminate signal and background electrons and $P_{(S,B),i}$ is the probability density function for the i^{th} input variable for signal (S, derived from $Z \to ee$ events) or background (B, derived from dijet events). Some of these input variables include hadronic leakage, primary track quality, cluster information from different parts of the ECAL, and track-cluster matching. $\mathcal{L}$ is parameterized as a function of η and E_T to maintain signal efficiency in different regions of the detector and for higher energy electrons, which deposit more of their energy in the deeper layers of the calorimeter. Additionally, $\mathcal{L}$ depends on the number of reconstructed vertices in the event, removing sensitivity to pile-up conditions. From simulated $Z \to ee$ and dijet events, the signal efficiency and corresponding background rejection are defined, allowing analyzers to

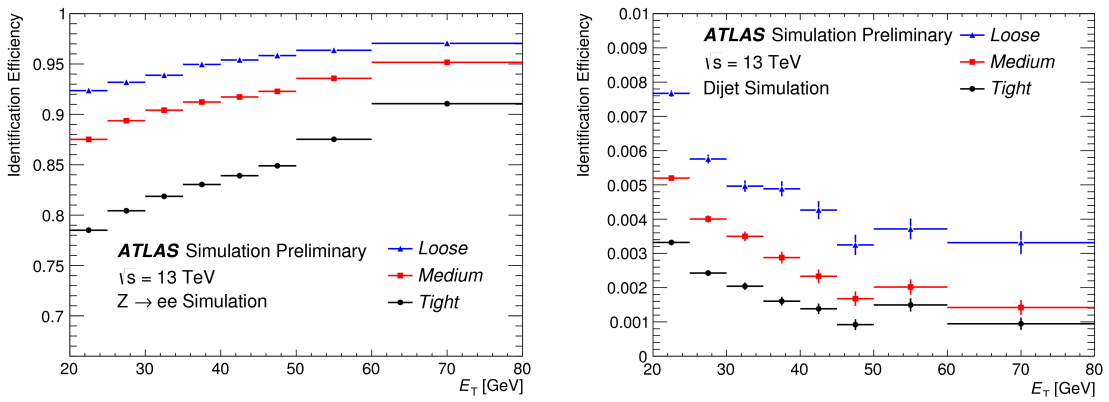

Figure 4.3: Identification efficiency in simulated $Z \to ee$ (left) and dijet (right) events for the recommended likelihood working points.

apply a selection on $\mathcal{L}$ to select electrons for analysis. Three levels are recommended for use in analyses: *loose*, *medium*, and *tight*. The relative identification efficiencies in simulation for each working point are shown in Fig. 4.3 as a function of E_T. The efficiencies for reconstruction and identification in both Data and MC are shown in Fig. 4.4. The difference between Data and MC results in a scale factor applied to simulated events to correct for the mis-modeling.

In addition to the identification criteria described above, electrons can also have an *isolation* requirement applied to reduce contamination from background electrons that may be non-prompt or mis-identified jets. Electron isolation is determined separately for the track and calorimeter components of the electron to determine the amount of activity near the electron. Calorimeter isolation, E_T^{cone}, first sums the transverse energies of clusters within a cone of radius $R = 0.2$ around the electron. The energy within a rectangle $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ centered on the electron barycenter is then subtracted. Finally, corrections for energy leakage, event-level pileup and underlying event are applied to determine a final isolation value for E_T^{cone}. Track isolation is based on the p_{T} of tracks that are not associated with, but are reconstructed near the electron. The isolation p_T^{cone} sums the p_{T} of the tracks within a cone of radius $R = \min(0.02, 10/E_T \text{ [GeV]})$ of the primary track. Similarly, $p_T^{\text{varcone}}(R_{\text{max}})$ uses a cone of radius $R = \min(10/p_{\text{T}} \text{ [GeV]}, R_{\text{max}})$ and the isolation must satisfy $p_T^{\text{varcone}}(R_{\text{max}}) < I p_{\text{T}}^{\ell}$, where I is a constant defined by the specific working point, and p_{T}^{ℓ} is the lepton p_{T}. Requirements on the track and calorimeter isolation, including $E_T^{\ell}-$ or $p_{\text{T}}^{\ell}-$dependent parameterizations, can increase the purity of signal electrons for analyses sensitive to non-prompt electrons.

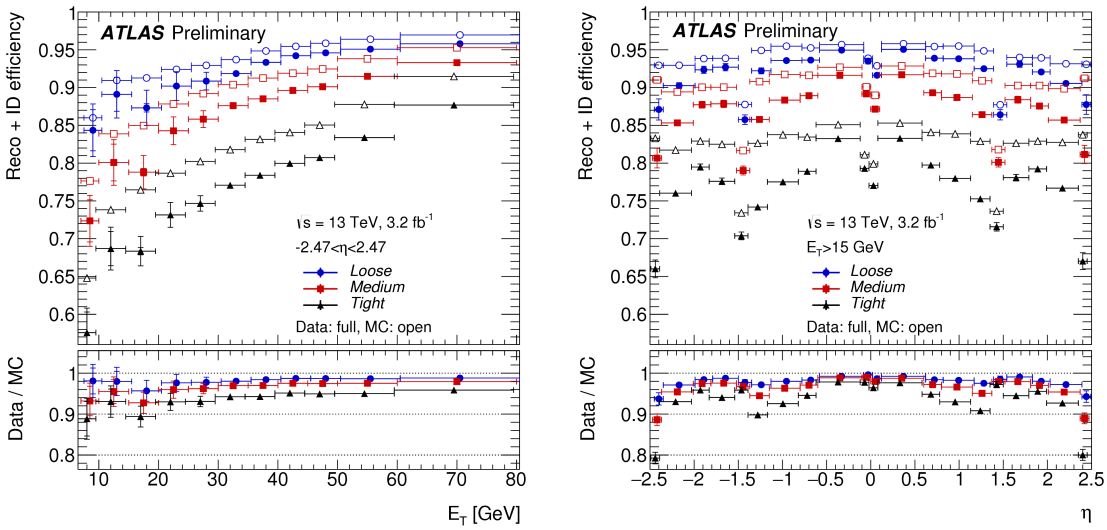

Figure 4.4: Reconstruction and identification efficiency of electrons in 2015 Data and MC events for each likelihood working point as a function of E_T (left) and η (right).

4.2.3 Muons

Muons are reconstructed using ID tracks with information from the MS in the central region, $|\eta| < 2.5$, of the ATLAS detector [61]. Muons are used throughout the analyses presented here to either select events and reconstruct the $T\bar{T}$ system, see Sect. 5.1, or as a veto to select fully hadronic final states, see Sect. 5.2. Within the MS, a search for detector hits in the MDT, CSC, RPC, and TGC first identifies *segments*. A fit of the segments across different layers results in the identification of muon track candidates. Segments found in the middle layers of the MS initially seed track candidates, which is extended to the innermost and outermost layers. In the track fitting procedure, criteria on the hit multiplicity and fit quality of segments improves the reconstruction. Additionally, at least two segments are required to build a track candidate except in the barrel-endcap overlap region, where a single high-quality segment with both η and ϕ information is acceptable to build a track. An overlap removal ensures that the same segment is not applied to multiple muon tracks unless two tracks share at most two out of three segments. Finally, a global χ^2 fit determines if a track candidate is accepted or rejected.

With the MS and ID track candidates identified, four types of muon reconstruction are performed: *combined* (CB) muons require a matching ID track and MS track, the combination of an ID track and only segments and no track in the MS results in *segment tagged* (ST) muons, *extrapolated* (ME) muons have no ID track to match the MS track (most common for $2.5 < |\eta| < 2.7$ where no ID tracks exist), and *calorimeter tagged* (CT) muons lack an MS

track and segments but have an ID track matched to calorimeter deposits consistent with a minimum-ionizing particle.

After muon candidates have been reconstructed and categorized as one of the four muon types described above, identification criteria improves the rejection against muons produced from hadron decays, e.g., pions or kaons, and maintains a high-efficiency for prompt muons. Muon identification criteria includes a measure of the consistency between q/p for muons in the ID and MS ('q/p significance'), the p_{T} balance of the MS and ID tracks relative to the combined p_{T} (ρ'), and the χ^2 of the combined track fit. Four identification working points exist for analyses to apply: loose, medium, tight, and high-p_{T}. The defined working points are inclusive, e.g., *medium* is a subset of *loose*, except for the high-p_{T} working point which has been defined primarily for analyses searching for high mass resonances. All four muon types are used for the loose identification. CT and ST muons are only included in the region $|\eta| < 0.1$ and ME muons are only used for the $|\eta| > 2.5$ regions. CB muons are selected in $|\eta| < 2.5$ with requirements on the number of hits in the MDT and q/p significance. Medium muons are a subset of the loose muons that only consist of the CB muons described in loose working point. Tight muons must also pass the medium selection with more stringent requirements on the number of hits in the MS and χ^2 fit. A two-dimensional selection on the q/p significance and ρ' as a function of muon p_{T} further improves rejection of background muons while reducing the signal efficiency, compared to the medium working point. High-p_{T} muons are identified from CB muons that pass the medium selection with strict requirements on the number hits in select regions of the MS. The muon p_{T} resolution is maximized for $p_{\mathrm{T}} > 1.5$ TeV relative to the other defined working points. Furthermore, the tracks associated to a selected μ must satisfy $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma_{d_0}| < 3$. The efficiency for different working points and as a function of p_{T} is provided in Fig. 4.5.

In addition, the relative isolation of muons with respect to other activity in the detector can also be used in muon identification and reconstruction. Through a harmonization effort within the ATLAS collaboration, the muon isolation definitions are identical to those described for electrons in Sect. 4.2.2.

4.2.4 Photons and Tau Leptons

Photons and τ leptons are not considered in the analyses presented in this dissertation. The following section contains a brief overview of photon [62] and τ lepton [63] identification and reconstruction.

Photon reconstruction uses a similar method as described for electrons in Sect. 4.2.2. Converted photons are identified by oppositely charged tracks that are associated with an ECAL cluster. The tracks associated with the ECAL cluster must also be collinear at the

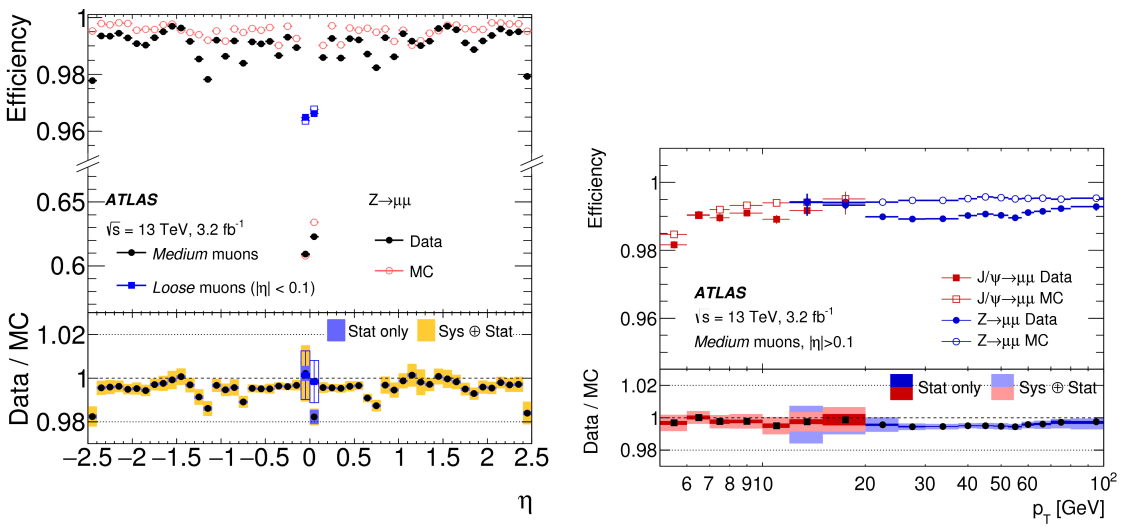

Figure 4.5: Reconstruction and identification efficiency of muons in 2015 Data and MC events as a function of η (left) and muon p_{T} (right).

production vertex and have properties consistent with electrons. If a single track is matched to an ECAL cluster but does not contain hits in the innermost layer of the pixel detector, it can be considered a converted photon rather than an electron. ECAL clusters without associated tracks are classified as unconverted photons. To first classify photon candidates as signal or background, isolation requirements are imposed on the ECAL cluster. The same isolation requirements definitions in Sect. 4.2.2 are applied, where the cone radius is fixed to 0.2 rather than p_{T} or E_T dependent, and a selection on the isolation quantities increases signal purity. Furthermore, non-prompt photons, e.g., hadronic decays, are distinguished from prompt photons using properties of the ECAL shower structure, including hadronic leakage, ratio of cell energies within ECAL layers, and width of the shower in different layers. The selection is optimized separately for converted and unconverted photons and in different bins of η to account for material changes in the detector.

Reconstruction of τ leptons primarily focuses on the hadronic decays of τ leptons. The leptonic decays result in electrons or muons with missing transverse momentum that can be identified using the methodology described in Sects. 4.2.2 and 4.2.3. Hadronic τ decays are classified based on the number of charged pions, 1– or 3–pronged (1 or 3 associated tracks), produced in the final state. These decays are reconstructed using the same algorithms as jet reconstruction, described in Sect. 4.2.5, as they rely upon energy deposits in the HCAL. The τ lepton vertex may not coincide with the designated primary vertex in the event, and thus the τ vertex is chosen as that which maximizes the sum of the ${p_{\mathrm{T}}}^2$ of tracks associated to the vertex candidate among all vertex candidates that are geometrically matched to a given

candidate cluster. Tracks are associated to a τ candidate if they are within a cone of radius 0.2 around the τ axis. The number of tracks assigned to a τ candidate are used to classify the τ decay topology. Finally, τ lepton identification relies on a multi-variate technique that uses information concerning τ candidate properties such as the energy distribution, track momenta, and angular separation of track candidates. Inputs to the multi-variate analysis are dependent on the classification of a τ candidate as either 1– or 3–pronged.

4.2.5 Jets

In the ATLAS detector, collimated sprays of particles that are produced from the hadronization of quarks are referred to as *jets*. Unlike the previous discussions of reconstructed physics objects, jets do not represent a particle in the SM. Instead, jets are used as a tool to represent quarks and gluons produced directly in the pp collisions or from hadronic decays of heavy objects like the electroweak or Higgs bosons, top quarks, or new physics. Combining information from the calorimeters and ID provides an estimation of the four-momentum of the particle that originated the shower. The choice of the jet definition, and assumptions therein, influences the analysis and interpretation of results. Jets are used throughout the analyses presented in Sects. 5.1 and 5.2 to identify the decays of heavy objects produced by VLQs and discriminate against SM backgrounds.

In the first step towards jet reconstruction, calorimeter cells are topologically combined to form *topo-clusters* [64]. Topo-cluster formation begins with determining the cell significance for each calorimeter cell. The cell significance is defined as the ratio of the energy measured in the cell $E^{\mathrm{EM}}_{\mathrm{cell}}$ and the average noise level $\sigma^{\mathrm{EM}}_{\mathrm{noise,cell}}$, see Eq. 4.2, where the cell energy is measured at the EM scale and the noise level is simulated using input from run 1 data with adjustments for different data-taking conditions.

$$\varsigma^{\mathrm{EM}}_{\mathrm{cell}} = \frac{E^{\mathrm{EM}}_{\mathrm{cell}}}{\sigma^{\mathrm{EM}}_{\mathrm{noise,cell}}} \tag{4.2}$$

Cells that have $\varsigma^{\mathrm{EM}}_{\mathrm{cell}} > S$, where S is optimized based on test-beam experiments, are referred to as seed cells. [1] Once all seed cells have been identified, the topological clustering begins. Cells adjacent to a seed cell, within the same layer or adjacent layers that overlap in (η,ϕ), are combined with the seed cells to form proto-clusters. If a non-seed cell in the proto-cluster satisfies $\varsigma^{\mathrm{EM}}_{\mathrm{cell}} > N$, for $N < S$, its neighboring cells are also combined into the proto-cluster. This procedure continues iteratively until the last cells combined into the proto-cluster do not satisfy $\varsigma^{\mathrm{EM}}_{\mathrm{cell}} > N$. The final structure of a topo-cluster consists of three

[1] For run 2, topo-clusters are forbidden from being seeded by the presampler layers. This restricts jet formation from low-energy particles that do not penetrate the calorimeters.

levels of cells, seed cells $\varsigma_{\text{cell}}^{\text{EM}} > S$, growth cells $\varsigma_{\text{cell}}^{\text{EM}} > N$, and boundary cells $\varsigma_{\text{cell}}^{\text{EM}} > 0$. The boundary cells are responsible for collecting any energy not contained in the seed and boundary cells. The cell significances are set to $S = 4$ and $N = 2$, where the values are derived from optimizations of the calorimeter response and the relative energy resolution for charged pions in test-beam experiments [65]. Lastly, topo-clusters that contain two or more local maxima, $E_{\text{cell}}^{\text{EM}} > 500$ MeV, are split in three dimensions along the boundary between the maxima to create separate topo-clusters. To construct the four-momentum of the topo-cluster, the total energy is the sum of the energy in all cells, the mass is set to zero, and the angles are set to the barycenter of the topo-cluster.

Once the topo-clusters have been defined, they are clustered together using the anti-k_t clustering algorithm [66,67]. The anti-k_t algorithm is a sequential combination algorithm that is infrared and collinear safe. ATLAS uses the open source `FastJet` software library [68] to facilitate interpretation and comparison of experimental results within the research community. Distance measures are defined for all topo-clusters, d_{iB}, and between all topo-clusters d_{ij} in Eq. 4.3.

$$d_{ij} = \min\left(p_{T,i}^{2n}, p_{T,j}^{2n}\right)\Delta_{ij}^2, \quad d_{iB} = p_{T,i}^{2n}R^2 \tag{4.3}$$

where R is the radius parameter, $\Delta_{ij}^2 \equiv (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and $n = -1$ for the anti-k_t algorithm. Jets with $R \leq 0.6$ are referred to as small-radius (small-R) jets and those with $R > 0.6$ are referred to as large-radius (large-R) jets. The clustering sequence begins by identifying the smallest value d_{ij} or d_{iB}. If $d_{ij} < d_{iB}$, topo-clusters i and j are combined. If $d_{iB} < d_{ij}$, the topo-cluster is removed from the clustering sequence and identified as a jet. This process continues until all topo-clusters have been identified as jets. The reconstructed jet clustering sequence uses topo-clusters as inputs, but any four-vector object can be passed to the clustering algorithm. Truth particle inputs result in *truth jets* and *track jets* are built from individual tracks.

The measured properties of jets do not correctly reflect the underlying process due to extra activity in the calorimeter and other detector effects. A jet energy calibration corrects for these effects [69]. Depending on the radius parameter of the jet, a different calibration is applied. The following sections detail the calibration procedure for small-R jets and then large-R jets.

Small-R jets are first corrected in p_{T} for pile-up based on the jet area. The p_{T} density for each jet is defined as p_{T}/A, where A is the jet area defined by ghost-association [70], where particles of negligible energy and momentum are uniformly placed throughout the (η,ϕ) space of the detector and included in the clustering procedure with topo-clusters. The median p_{T} density of all jets is then used to correct the jet p_{T} and is shown for select values of

N_{PV} and μ in Fig. 4.6. Additional corrections for the number of primary vertices and average number of pp collisions per bunch crossing, which were found to be linearly independent in run 1, yield the final reconstructed jet p_{T}. Once the jet p_{T} has been corrected, jets undergo a simulation-based energy and η calibration. The average energy response $E^{\mathrm{reco}}/E^{\mathrm{truth}}$ for geometrically matched reconstructed and truth jets is used to correct the jet energy as a function of the detector position η_{det}. Fig. 4.7 shows the energy response for different truth energy values as a function of η_{det}.

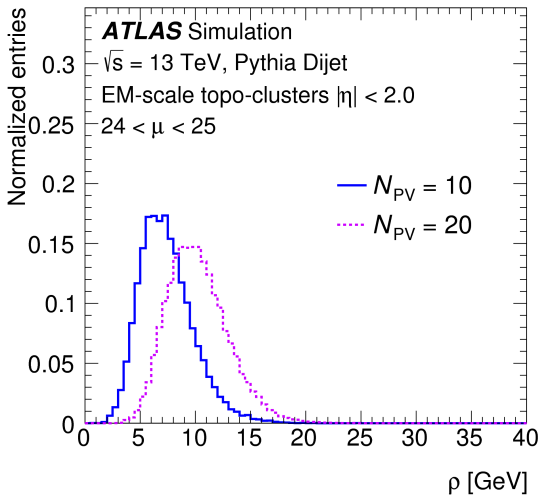

Figure 4.6: Median p_{T} density for jets under two N_{PV} scenarios and a single bin of μ.

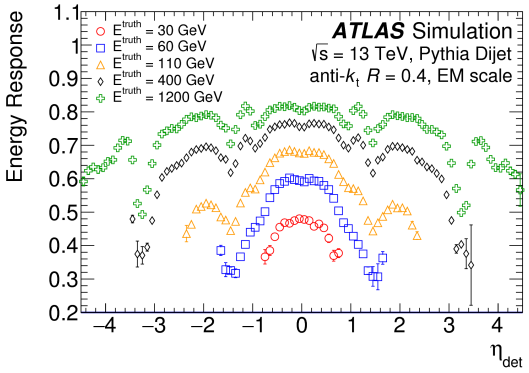

Figure 4.7: Jet energy response for different truth energy values as a function of η_{det}.

After these calibrations have been applied, a residual difference in properties of the transverse and longitudinal components of the jets is still observed. A *global sequential calibration* is applied to five observables, excluding correlations between them, to correct for properties of jets, namely the particle composition and shower shape. Finally, the difference

between data and MC prediction is corrected for events where jets are balanced against a well-known physics object, e.g., photon or Z boson.

Jets from the hard scatter are differentiated from jets produced by pile-up events using the Jet Vertex Tagger (JVT) algorithm [71]. The JVT discriminant is constructed as a two-dimensional likelihood in the corrJVF$-R_{\mathrm{pT}}$ plane. The corrJVF matches the tracks associated to each jet with different vertices in the collision, independent of pile-up in the event. R_{pT} is defined as the scalar sum of the p_{T} of all tracks associated to the jet, normalized to the total jet p_{T}, that are matched to the primary vertex. The classification of the nearest neighbors to the jet in the corrJVF$-R_{\mathrm{pT}}$ determines the likelihood that a given jet will be classified as originating from the hard-scatter or pile-up.

Topo-clusters are classified as being electromagnetic or hadronic using shower shapes and energy densities using local cluster weighting method. A calibration dependent on the classification properly corrects for detector effects that result in a mis-measurement of jet energy. Large-R jets are then built using the calibrated topo-clusters as inputs to the anti-k_t algorithm with $R = 1.0$.

After large-R jets are formed, a trimming procedure reduces sensitivity to pile-up and other soft energy that contaminates the jet [72]. To trim the large-R jets, schematically represented in Fig. 4.8, the jet constituents are clustered into $R = 0.2$ subjets using the k_t algorithm [73, 74], $n = 1$ in Eq. 4.3. Any subjet that has a p_{T} below $f_{\mathrm{cut}} = 5\%$ of the total large-R jet p_{T} is removed. The subjet radius parameter and trimming threshold f_{cut} were optimized using simulated W/Z boson and dijet events [75].

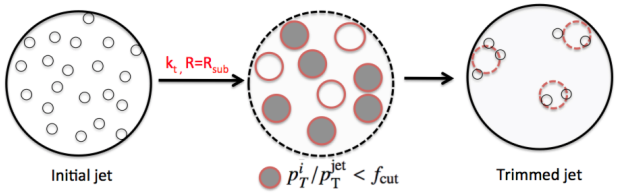

Figure 4.8: Jet trimming procedure for large-R jets. In ATLAS, $R_{\mathrm{sub}} = 0.2$ and f_{cut}=5%.

The jet energy and mass of trimmed large-R jets are calibrated using a similar treatment as small-R jets [69]. Corrections pertaining to pile-up, number of primary vertices, or interactions per bunch crossing are not applied to large-R jets because the trimming procedure removes such dependencies on the jet energy and mass. In addition to the energy calibration, the large-R mass is also calibrated using a MC-based technique. The jet mass response $m_{\mathrm{reco}}/m_{\mathrm{truth}}$ is determined from reconstructed and truth-level jets that are geometrically matched in the detector. The offset of the mean of the Gaussian fit to the jet mass response

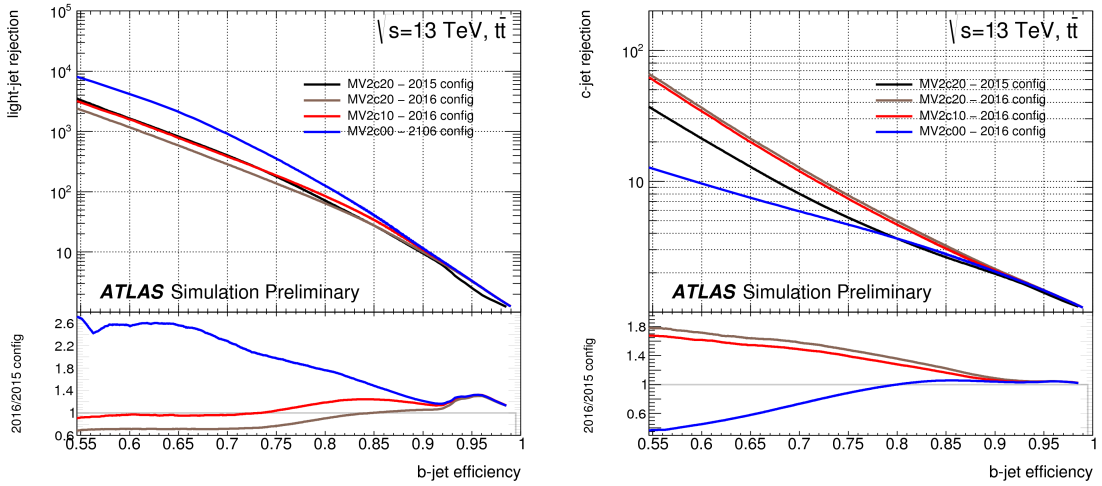

Figure 4.9: Performance of b-tagging algorithms shown for light-jet rejection (left) and c-jet rejection (right) as a function of signal efficiency.

from one is then used to correct the large-R jet mass, parameterized as a function of the jet m/p_{T}.

4.2.5.1 b-Tagging

In the ATLAS detector, b-quarks will hadronize into B-hadrons and travel distances $\mathcal{O}(mm)$ before decaying. The long lifetime results in displaced vertices and larger impact parameters associated with the B-hadron that can be used to tag b-quarks. A small-R jet is identified as originating from the hadronization of a b quark, referred to as b-tagged, using the `mv2c10` discriminant [76]. A multivariate approach to identify b-quarks results in the best signal efficiency and background rejection. The discriminant `mv2c10` uses information from the jet kinematics, impact parameters, secondary vertex parameters, and topology of the hypothesized decay chain [77]. Simulated events composed of with 90% light flavor quarks and 10% c-quark decays are used to maximize background rejection. Events from simulated $t\bar{t}$ events provide signal jets in the training procedure. The performance of the b-tagging algorithms are shown in Fig. 4.9. VLQ decays are expected to produce real b-quarks directly, e.g., $T \to Wb$, or indirectly, e.g., $T \to Ht \to (bb)(Wb)$. The b-tagging of small-R jets improves background rejection of processes that do not produce many real b-quarks, e.g, multi-jet, and enhances the sensitivity to new physics.

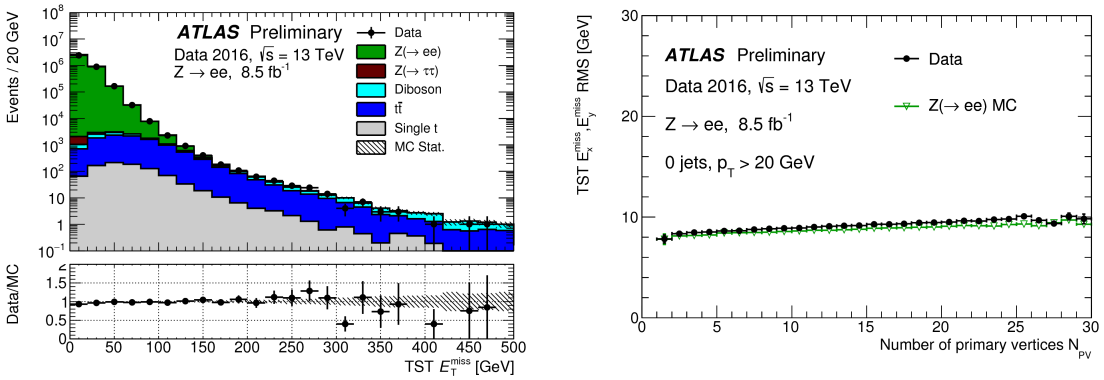

Figure 4.10: Comparison between Data and MC for TST $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and the $E_{\mathrm{T}}^{\mathrm{miss}}$ RMS as a function of the number of primary vertices (right).

4.2.6 Missing transverse momentum

For a given reconstructed event, the magnitude of the negative vector sum of the p_{T} of all reconstructed and calibrated objects is defined as the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$ [78].

$$\begin{aligned} E_{\mathrm{T}}^{\mathrm{miss}} &= \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \\ \phi^{\mathrm{miss}} &= \arctan\left(E_y^{\mathrm{miss}}/E_x^{\mathrm{miss}}\right) \\ E_i^{\mathrm{miss}} &= E_i^{\mathrm{miss},e} + E_i^{\mathrm{miss,jets}} + E_i^{\mathrm{miss},\mu} + E_i^{\mathrm{miss,soft}}, \quad i = x, y \end{aligned} \tag{4.4}$$

An extra term, added to account for *soft* energy in the event that is not associated to any of the selected objects, is calculated from ID tracks matched to the primary vertex. The track-based soft term (TST) reduces the $E_{\mathrm{T}}^{\mathrm{miss}}$ dependence on pile-up. The performance of $E_{\mathrm{T}}^{\mathrm{miss}}$ in early 2016 data are shown in Fig. 4.10. Good agreement between data and prediction is achieved across a wide range of $E_{\mathrm{T}}^{\mathrm{miss}}$ and number of primary vertices.

In the analyses presented in this dissertation, $E_{\mathrm{T}}^{\mathrm{miss}}$ is used to reject multi-jet events, see Sect. 5.1, and in the data-driven background estimation for multi-jet events, see Sect. 5.2. Additionally, the analysis described in Sect. 5.2 uses $E_{\mathrm{T}}^{\mathrm{miss}}$ to maintain orthogonality with a separate VLQ search at ATLAS [2].

CHAPTER V

Searches for Vector-like Quarks

The following two sections describe separate, but similar, analyses searching for pair production of VLQs. Additionally, the analyses are orthogonal by construction in the number of leptons in the final state. The first analysis, $T\bar{T} \to Wb + X$ [48], uses a dataset collected in 2015 and 2016 with an integrated luminosity of 14.7 fb^{-1}. The $T \to Wb$ decay mode is targeted using the one lepton final state. The second analysis, $T\bar{T}/B\bar{B} \to$ jets, uses the full dataset collected in 2015 and 2016 with an integrated luminosity of 36.1 fb^{-1}. The fully hadronic final state exploits that large branching ratio of boson decays to quarks and complements the existing one and multi-lepton final state searches in ATLAS. In the following sections, the mass of the VLQ at *truth-level*, i.e., generated by simulation, is written as m_T, m_B, or generically as m_{VLQ}.

5.1 $T\bar{T} \to Wb + X$

The analysis strategy is optimized for a signal hypothesis of two top-like VLQs that decay to $T\bar{T} \to Wb + X$, where $X = Wb, Zt, Ht$. In this signal scenario, single lepton final states have strong sensitivity due to the relatively large branching ratio and clean signature against the multi-jet background in pp collisions. Among the other 1-lepton analyses searching for VLQs, $Ht + X$ [2] and $Zt + X$ [3], the final state $Wb + X$ was selected because of the dominant branching ratio in the singlet scenario, highlighted in Fig. 2.4. The representation from Protos assumes a branching ratio of 0 in the doublet scenario, but this is not an explicit constraint on all models that include VLQs, and some sensitivity to doublet models is expected.

The lepton, e or μ, is assumed to have been produced by the W boson from the T decay and the events are identified with the un-prescaled single lepton triggers. The dominant SM $t\bar{t}$ background is significantly reduced by splitting the analysis into orthogonal boosted and resolved channels. The requirement of a large-R jet identified as a W boson distinguishes

the boosted and resolved scenarios. Non-$t\bar{t}$ backgrounds, such as W+jets and Z+jets, are reduced by requiring multiple b-tagged small-R jets. The neutrino is reconstructed using the lepton, $E_{\mathrm{T}}^{\mathrm{miss}}$, and assuming the W mass constraint. Events are categorized based on the scalar sum of the transverse momentum of all reconstructed objects S_{T} and the angular separation between the lepton and the neutrino $\Delta R(\ell, \nu)$. Finally, a profile likelihood fit is performed using the reconstructed VLQ mass to extract the excluded cross section as a function of the T mass and branching ratios. In this section, the *reconstructed* leptonic (hadronic) VLQ mass is written as m_T^{lep} (m_T^{had}) to distinguish this quantity from the truth-level VLQ mass, m_T.

5.1.1 Data and Simulated Samples

The analyzed dataset combines the full recorded data from 2015 with the data collected until the end of August 2016 with the ATLAS experiment. This corresponds to an integrated luminosity of 14.7±0.4 fb^{-1}. Only data collected when all detector components were fully operational are considered for this analysis.

5.1.1.1 Signal production

Simulated $T\bar{T}$ events are generated with the leading-order (LO) generator Protos v2.2 [47] using the NNPDF2.3 LO PDF set and passed to Pythia 8.186 [79] for parton showering and fragmentation. The A14 [80] set of optimized parameters for the underlying event description using the NNPDF2.3LO PDF set is used. VLQs are produced in the $SU(2)$ singlet model, but are forced to decay with a branching ratio of 1/3 to each of the three modes (Wb, Zt, and Ht). Arbitrary sets of branching ratios consistent with the three decay modes summing to unity are obtained by re-weighting the samples using particle-level information. This is accomplished by weighting events by $(\mathrm{BR}(T \to A) \times \mathrm{BR}(\bar{T} \to B))/(1/9)$, where $T \to A$ and $\bar{T} \to B$ (A,$B = Wb$, Zt, and Ht) are the actual decay modes that took place in the event, and $\mathrm{BR}(T \to X)$, $\mathrm{BR}(\bar{T} \to Y)$ are the desired branching ratios. The branching ratios for a given model vary as a function of VLQ mass, as shown in Fig. 2.4 for three possible models. Fig. 2.4 also shows the increase in VLQ production cross sections due to the increase in centre-of-mass energy in run 2. In all simulated samples used in this search, the top quark and SM Higgs boson masses are set to 172.5 GeV and 125 GeV, respectively.

VLQ samples are produced for hypothetical masses of 500, 600, 700-1200 (steps of 50 GeV), 1300, and 1400 GeV. The pair production cross sections vary from 3.38 ± 0.25 pb ($m_T = 500$ GeV) to 3.50 ± 0.43 fb ($m_T = 1400$ GeV), computed using `top++` v2.0 [44] at next-to-next-to-leading order (NNLO) in QCD, shown in Fig. 2.3. The calculations include

resummation of next-to-next-to-leading logarithmic soft gluon terms, and using the MSTW 2008 NNLO set of parton distribution functions (PDF). Theoretical uncertainties result from variations on the factorization and renormalization scales, as well as from uncertainties on the PDF and α_S. The latter two represent the largest contributions to the overall theoretical uncertainty on the cross section and were calculated using the PDF4LHC [81] prescription with the MSTW 2008 68% CL NNLO, CT10 NNLO [82, 83] and NNPDF2.3 5f FFN PDF sets [84].

5.1.1.2 Background production

The main SM backgrounds considered in this analysis are $t\bar{t}$, W+jets, Z+jets, diboson, single top, and $t\bar{t}+V$. QCD multi-jet events also have a small contribution to the analysis, but these contributions are estimated using data-driven methods, as described in Sect. 5.1.4.

The $t\bar{t}$ MC samples are generated by POWHEG-BOX +PYTHIA 6 [85, 86], with a one-lepton filter. using the `Perugia2012 tune` [87] for the showering and the CT10 PDF set, and setting the hdamp parameter to the mass of the top-quark. Additional samples are used to evaluate modeling uncertainties on the $t\bar{t}$ sample, which are generated using POWHEG-BOX +HERWIG ++ [88] and `aMC@NLO`+HERWIG ++. Further, POWHEG-BOX +PYTHIA 6 samples are generated varying the factorization and normalization scale by 2 and 0.5, as well as the NLO radiation factor, hdamp, between m_{top} (172.5 GeV) and twice m_{top}.

Single top production ($Wt-$, $s-$, and $t-$channel) is also generated by POWHEG-BOX +PYTHIA 6, using the `Perugia2012 tune` for the showering and the CT10 PDF set.

The W+jets, Z+jets, and diboson (WW,WZ,ZZ) samples use the SHERPA v2.2 generator [89] with the CT10 PDF set. To improve the number of MC statistics available for analysis, the W/Z+jets samples are produced in slices of $p_{\text{T}}(W/Z)$ with filters on the heavy-flavor content (light-quarks, c-quarks, b-quarks).

The $t\bar{t}+V$ background is modeled using samples produced with MADGRAPH +PYTHIA 8 [90]. The contribution of $t\bar{t}+V$ events to the event selection is very small.

The MC samples are based on a full `GEANT4` simulation of the ATLAS detector [55, 91] and are reconstructed using the same analysis chain as the data, described in Sect. IV.

5.1.2 Event Selection

Selections on the simulated samples and data are applied to enhance the expected signal acceptance to background rejection, and thus the overall sensitivity to discovery. These selections are applied in multiple steps to ensure a thorough understanding of all SM processes that produce final states similar to that of our expected signal. In this analysis, the event

selection is optimized using MC samples described in Sect. 5.1.1. The final signal region is defined after performing the full statistical analysis, including systematic uncertainties, over different options using MC-only and maximizing the expected limit on the VLQ mass with $\mathrm{BR}(T \to Wb) = 100\%$.

Initially, a baseline is chosen such that all backgrounds can be thoroughly examined and software frameworks can be validated across different analyses (most analyses use a common analysis framework with custom extensions relevant for their analysis that need to be cross-checked). Events must pass one of the unprescaled single lepton (e or μ) triggers listed in Table 5.1. A logical `OR` of all triggers is applied for selecting the events. Multiple triggers differ in their relative p_{T} thresholds and identification criteria to recover events with leptons in different topologies and kinematic regions.

Table 5.1: Single lepton triggers applied for the $T\bar{T} \to Wb + X$ analysis. There are separate triggers for 2015 and 2016 due to the different LHC pp collision configurations, e.g., instantaneous luminosity. These triggers are common among single lepton analyses in the ATLAS Top Working Group.

Year	Lepton Flavor	HLT Triggers (logical `OR`)
2015	Electron	`e24_lhmedium_L1EM20VH`, `e60_lhmedium`, `e120_lhloose`
2016	Electron	`e24_lhtight_nod0_ivarloose`, `e60_lhmedium_nod0`, `e140_lhloose_nod0`
2015	Muon	`mu20_iloose_L1MU15`, `mu50`
2016	Muon	`mu26_ivarmedium`, `mu50`

After an event is selected by the triggers, the events must pass requirements on the data quality and the existence of at least one primary vertex (see Sect. 4.2.1). Next, the physics objects are reconstructed according to the working points described in Sect. 4.2. In this analysis, electrons are used if they satisfy the tight likelihood working point, a p_{T}-dependent isolation ($\equiv$'gradient isolation') that maintains an efficiency of $\epsilon = (0.1143 \times p_{\mathrm{T}}[\text{ GeV}] + 92.14)\%$, and a p_{T} above 25 GeV. Muons are identified with the medium working point, the gradient isolation, and a minimum p_{T} of 25 GeV. Small-R jets must have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. Selected jets that are centrally-located ($|\eta| < 2.4$) with low-p_{T} ($p_{\mathrm{T}} < 60$ GeV) must satisfy JVT> 0.59. A small-R jet is considered b-tagged if it satisfies the 77% efficiency working point of the `mv2c10` algorithm. The trimmed large-R jets are only required to pass a simple kinematic selection of $p_{\mathrm{T}} > 200$ GeV, mass larger than 50 GeV, and $|\eta| < 2.0$.

The different kinds of physics objects are reconstructed independently and it is possible for two different objects, e.g., electrons and jets, to duplicate detector information and result

in double-counting energy. To prevent these issues, an overlap removal procedure is applied to leptons and jets using their relative separation in the detector $\Delta R(\ell, \text{jet})$. If the selected electron is reconstructed within $\Delta R(e, \text{jet}) < 0.2$ of a given jet, the jet is removed from the event. However, if the electron is reconstructed within $0.2 < \Delta R(e, \text{jet}) < 0.4$ of a jet, the electron is removed from the event. A p_{T}-dependent overlap removal working point is applied to muons and jets to recover high–p_{T} muons produced near hadronic activity. Muons are removed if $\Delta R(\mu, \text{jet}) < \min\left(0.4, 0.04 + 10/p_{\text{T}}\ [\text{GeV}]\right)$ and jets are removed if one of the two following conditions is met:

1. The jet has less than 3 tracks ($p_{\text{T}} > 500$ MeV) associated to the primary vertex or ($p_{\text{T}}^{\mu}/\texttt{sumTrkPt500} > 0.7$ and $p_{\text{T}}^{\mu}/p_{\text{T}}^{\text{jet}} > 0.5$)

2. The muon ID track is ghost associated to the jet or $\Delta R(\mu, \text{jet}) < 0.2$

where `sumTrkPt500`is the sum of the p_{T} of all tracks with $p_{\text{T}} > 500$ MeV. No overlap removal procedure is applied between large-R jets and other objects in the detector.

The final step of the baseline selection requires exactly one lepton (electron or muon) also identified by the trigger, and if the event contains three small-R jets, an additional large-R jet is required, otherwise only four small-R jets are required. The final states are referred to as e+jets (μ+jets) if the lepton is an electron (muon) and ℓ + jets if no distinction between lepton flavor is made.

Data and MC comparisons in a $t\bar{t}$–rich final state of at least four jets and two b-tagged jets are shown in Fig. 5.1. These distributions are used to compare the data and prediction with systematic uncertainties for the lepton p_{T}, transverse mass of the W bosons[1], p_{T} of the leading b-tagged jet, and $E_{\text{T}}^{\text{miss}}$. The green bands include all detector-related and $t\bar{t}$ modeling uncertainties described in Sect. 5.1.5.

After the baseline selection, the neutrino, ν, is reconstructed using the $E_{\text{T}}^{\text{miss}}$ and identified lepton, ℓ, by requiring the lepton-neutrino system to have an invariant mass equal to $m_W = 80.4$GeV. From this, the neutrino four-vector can be identified up to a two-fold ambiguity in p_z^{ν}:

$$p_z^{\nu} = \frac{\mu^2 p_z^{\ell}}{(E^{\ell})^2 - (p_z^{\nu})^2} \pm \sqrt{\frac{\mu^4 (p_z^{\ell})^2}{((E^{\ell})^2 - (p_z^{\nu})^2)^2} - \frac{(E^{\ell})^2 (E_{\text{T}}^{\text{miss}})^2 - \mu^2}{(E^{\ell})^2 - (p_z^{\nu})^4}} \tag{5.1}$$

where $\mu^2 \equiv \frac{1}{2} m_W^2 + p_x^{\nu} p_x^{\ell} + p_y^{\nu} p_y^{\ell}$. In the case of two real solutions, the one with a smaller absolute value is used. However, many events in the relevant phase space are found to

[1]The transverse mass of the W is defined as $m_{\text{T}}^W \equiv \sqrt{2 p_{\text{T}}^{\ell} E_{\text{T}}^{\text{miss}} (1 - \cos \Delta\phi(\ell, E_{\text{T}}^{\text{miss}}))}$

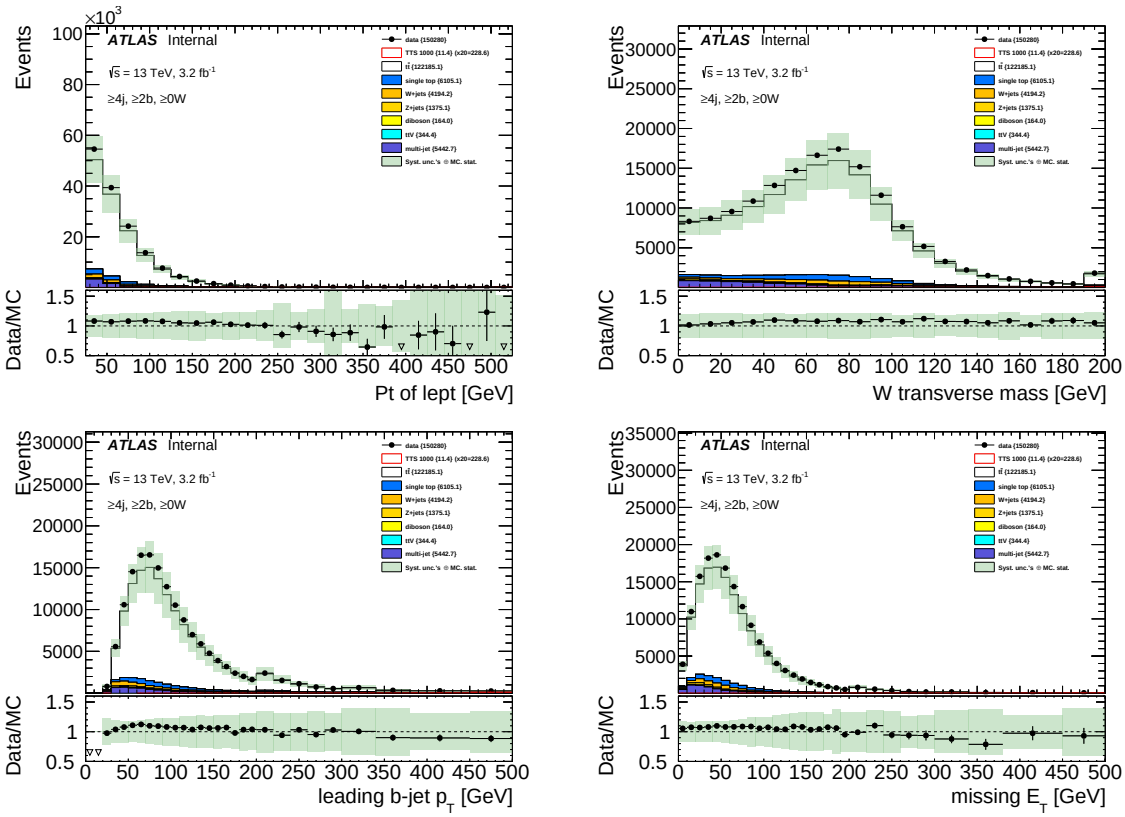

Figure 5.1: Data to MC comparison plots of ℓ + jets events for the lepton p_{T} (top left), transverse mass of the W boson (top right), leading b-tagged jet p_{T} (bottom left), and $E_{\mathrm{T}}^{\mathrm{miss}}$ (bottom right). The green band includes MC statistics, all $t\bar{t}$ modeling, and detector-related systematic uncertainties.

produce two imaginary solutions. The challenge is thus to determine a real solution that most closely reproduces the truth neutrino. The following algorithms were tested:

- "Real Only": Take only the real part of the solution.
- "Co-linear": Assume the neutrino is co-linear with the charged lepton, $\eta^{\nu} = \eta^{\ell}$ and $\phi^{\nu} = \phi^{\ell}$.
- "modCo-linear": Assume the neutrino has the same η as the lepton $\eta^{\nu} = \eta^{\ell}$ and $\phi^{\nu} = \phi^{\mathrm{miss}}$
- "Rotation": Rotate ϕ^{miss} until the solution is real.
- "scaleMET": Scale $E_{\mathrm{T}}^{\mathrm{miss}}$ until the solution is real. $E_{\mathrm{T}}^{\mathrm{miss}}$
- "Default": Use `TMinuit` to find the p_x^{ν} and p_y^{ν} that result in a real solution to the mass requirement with the minimum change in p_{T}^{ν}.

- "TMinuit": A Python implementation of "Default" used for validation.

Fig. 5.2 shows the agreement between the truth and reconstructed neutrino p_z and p_{T} for each method. It was found the "Default" reconstruction method best reproduces the true neutrino p_{T} and p_z.

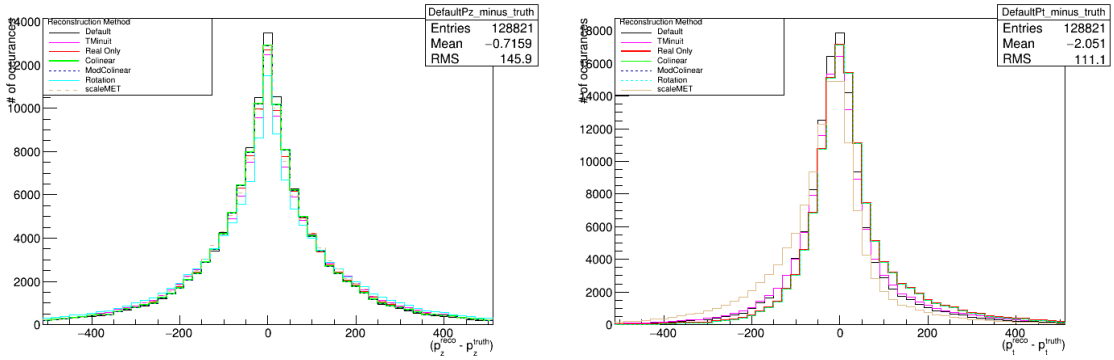

Figure 5.2: Comparison of different neutrino reconstruction methods to the truth values for neutrino p_z (left) and p_{T} (right).

To further improve the signal to background ratio, additional selection is required. In this next phase, the selection is optimized assuming a unity branching ratio BR($T \to Wb$) = 1. This particular selection is referred to as the *loose* selection.

The selection is motivated from considerations of the signal and background topologies and kinematics. In the $T\bar{T}$ decays, real b-quarks are produced, unlike SM processes involving light-flavor jets, and expected to be reconstructed as b-tagged small-R jets. Decays of VLQs will deposit large amounts of energy in the detector, $S_{\mathrm{T}} \sim 2m_{\mathrm{VLQ}}$, where S_{T} is defined as the scalar sum of all small-R jet p_{T}, lepton p_{T}, and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event. Signal-like events are assumed to have a topology consistent with a leptonically decaying W boson, where the lepton and neutrino will be reconstructed relatively close together and the high-p_{T} neutrino will be responsible for the $E_{\mathrm{T}}^{\mathrm{miss}}$. This selection also identifies two kinematic regimes for consideration, *boosted* and *resolved*, distinguished by the treatment of the hadronically decaying W boson, $W_{\mathrm{had}} \equiv W \to q\bar{q}'$. The following cuts are applied to select events for further analysis:

- $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV
- At least one W_{had} candidate
- At least one b-tagged small-R jet
- $\Delta R(\ell, \nu) < 0.8$

- S_{T}> 1200 GeV

The boosted region targets VLQ masses $\gtrsim 1$ TeV where the decay products are created with $p_{\mathrm{T}} > 200$ GeV in more than 90% of events and W_{had} can not be efficiently reconstructed using small-R jets. Instead, large-R jets are used to capture the decay products of W_{had}. A key distinction between SM $t\bar{t} \to WbWb$ and BSM $T\bar{T} \to WbWb$ is the mass of the particles producing the W bosons and b-quarks. Top quarks can be produced with high–p_{T} and the decay products (W,b) merge into a single large-R jet. T quarks are produced with low transverse momentum and decay to boosted W bosons and b-quarks that are well-separated in the detector $\Delta R(W, b) \sim \pi$. Large-R jets are identified as W bosons using a custom implementation of the 50% signal efficiency W and top taggers provided by the ATLAS collaboration [75, 92]. A combination of substructure variables, $D_2^{\beta=1}$ [93, 94], τ_{32}^{wta} [95,96], and mass are used to distinguish between jets that are W bosons and top quarks. The substructure variable $D_2^{\beta=1}$ is defined such that two-pronged decays are distinguished between one-pronged decays.

$$D_2^{\beta=1} \equiv \frac{e_3^{\beta=1}}{(e_2^{\beta=1})^3} \tag{5.2}$$

where e_i^{β} are the energy correlation functions:

$$\begin{aligned} e_2^{\beta} &= \frac{\mathrm{ECF2}(\beta)}{\mathrm{ECF1}(\beta)^2} = \frac{\sum\limits_{i<j\in J} p_{\mathrm{T},i} p_{\mathrm{T},j} \left(\Delta R_{ij}\right)^{\beta}}{\left(\sum\limits_{i\in J} p_{\mathrm{T},i}\right)^2} \\ e_3^{\beta} &= \frac{\mathrm{ECF3}(\beta)}{\mathrm{ECF1}(\beta)^3} = \frac{\sum\limits_{i<j<k\in J} p_{\mathrm{T},i} p_{\mathrm{T},j} p_{\mathrm{T},k} \left(\Delta R_{ij} \Delta R_{ik} \Delta R_{jk}\right)^{\beta}}{\left(\sum\limits_{i\in J} p_{\mathrm{T},i}\right)^3} \end{aligned} \tag{5.3}$$

The substructure variable τ_{32}^{wta}, is defined as

$$\tau_N = \frac{1}{\sum\limits_k p_{\mathrm{T},k} R} \sum_k p_{\mathrm{T},k} \times \min(\delta R_{1k}, \delta R_{2k}, ..., \delta R_{Nk}) \tag{5.4}$$

where the sum runs over the constituents k and δR_{ik} is the distance between subjet i and k. In the "winner-take-all" (wta) scheme, the direction of the highest–p_{T} topo-cluster within each k_T ($R = 0.2$) subjet of the large-R jet, defined from the trimming procedure, is used in the distance calculation rather than the direction of the subjet. Both the W– and top–tagging methods maintain a constant tagging efficiency as a function of the large-R jet p_{T}. The W tagger places a large-R jet into one of four categories depending on the mass and

substructure: fails identification (category 0), passes only $D_2^{\beta=1}$ (category 1), passes only mass (category 2), or passes both $D_2^{\beta=1}$ and jet mass (category 3). The top tagger simply determines if a jet passes or fails the requirements on τ_{32}^{wta} and jet mass. For the boosted region, the large-R jet is considered tagged if it falls into the W tagging categories 1, 2, or 3 and fails the top tagging identification. The top-tag veto significantly reduces the dominant SM $t\bar{t}$ background relative to the $T\bar{T}$ signal. If more than one large-R jet satisfies these requirements, the jet with mass closest to the W boson is selected.

In the $T\bar{T}$ topology it is possible for objects from the leptonically decaying T to be reconstructed near the large-R jet (e.g., the b-quark from the leptonically decaying VLQ may be reconstructed within the same large-R jet as W_{had}). In such instances, it becomes possible to mis-identify W_{had} because the large-R jet no longer maintains the properties of an isolated W boson. The standard tagger performance suffers in VLQ events because the large-R jets contain energy from objects other than W_{had}. This quantify the impact of this effect, large-R jets from VLQ events are selected if they are reconstructed close to the truth W that decays hadronically, $\Delta R(\text{large-}R \text{ jet}, W_{\text{had}}^{\text{truth}}) < 0.75$ and studied as a function of the separation between the truth hadronically-decaying W, $W_{\text{had}}^{\text{truth}}$, and the truth b-quark originating from the leptonically decaying VLQ, $b_{\text{lep}}^{\text{truth}}$, $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}})$. Both of the large-$R$ jet substructure variables $D_2^{\beta=1}$ and τ_{32}^{wta} have different trends for $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}}) < 1.0$ compared to $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}}) > 1.0$. When the $b_{\text{lep}}^{\text{truth}}$ overlaps with the $W_{\text{had}}^{\text{truth}}$, the energy from each object is contained in the same large-R jet and this results in a more three-pronged than two-pronged jet substructure ($\tau_{32}^{\text{wta}} \to 0.0$) and $D_2^{\beta=1}$ identifies that the jet is no longer two-pronged. Additionally, the mass for large-R jets is vastly different in events with $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}}) < 1.0$. For separations beyond 1.0, the $b_{\text{lep}}^{\text{truth}}$ is typically reconstructed as a small-R jet outside the large-R jet and the distributions for all substructure variables is constant, as shown in Fig. 5.3 for $T\bar{T}$ ($m_T = 1$ TeV) events. The efficiencies for tagging large-R jets as a function of $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}})$ are shown in Fig. 5.1.2. The efficiencies for the different categories for the W-tagger are plotted separately (=1, =2, =3), together (> 0), and alongside the top tagger efficiency. A similar trend as shown in the substructure plots exists in the tagging efficiency curves. For $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}}) < 1.0$ the efficiency for tagging the large-R jet as a W boson in categories 2 or 3, where the large-R jet mass is high, is approximately 0. The jet W-tagging efficiency for substructure only ($= 1$) and top tagger perform well in this regime as the large-R jet has a high mass (near the top quark) and a three-pronged structure. The W-tagger returns to its nominal efficiency of 50% and the top-tagger has a very low efficiency (<15%) for $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}}) > 1.0$, indicating the W bosons produced from VLQ decays are well-isolated. If there are multiple large-R jets passing the

above criteria, the jet with a mass closest to the W boson mass is identified as W_{had} of the event.

The resolved region complements the boosted region by targeting lower mass $T\bar{T}$ events where two small-R jets are used to identify W_{had}. This region also recovers events missed in the boosted region due to the large-R jet tagging requirements. At least four small-R jets and no large-R jet that passes the boosted criteria are required. Two small-R jets, j_1 and j_2, that satisfy $\Delta R(j_1, j_2) < 1.2$ are used to identify W_{had}. If there are multiple pairs of small-R jets passing the above selection, the pair (j_1, j_2) with invariant mass closest to the W mass is chosen as the W_{had} of the event. Either j_1 or j_2, but not both, can be b-tagged.

Table 5.2: Event yields for background and an example signal point after the boosted and resolved signal selections. The numbers correspond to an integrated luminosity of 14.7fb^{-1}. The quoted uncertainties include statistical and systematic components.

Sample	Boosted yields	Resolved yields
$t\bar{t}$	2604.63 ± 1013.98	3164.76 ± 1349.33
W+jets	470.55 ± 116.66	614.42 ± 167.28
Single top	191.00 ± 23.38	205.86 ± 30.91
Diboson	62.74 ± 81.85	54.02 ± 16.10
$t\bar{t} + V$	33.98 ± 3.87	33.63 ± 4.54
Multi-jet	80 ± 80	100 ± 100
Z+jets	20.83 ± 6.68	27.80 ± 7.65
Total Background	3463.73 ± 1009.26	4100.49 ± 1352.18
Signal (m_T =800 GeV, BR($T \to Wb$)=1)	286.45±19.05	117.17±11.87

Distributions at each step of the boosted and resolved selection are shown in Fig. 5.4 and Fig. 5.5, respectively. In these figures, each variable in the boosted event selection is shown just before a selection on that variable is applied, but with all previous cuts applied. The uncertainty bands includes statistical and $t\bar{t}$ modeling components. The event yields after the loose selection for background processes and $T\bar{T}$ (m_T = 800 GeV, $B(T \to Wb)$=1) are listed in Table 5.2.

After the final selection, a full reconstruction of the $T\bar{T}$ system is achieved. The leptonic W (lepton and neutrino) and hadronic W (W_{had}) are combined with small-R jets to build the $T\bar{T}$ candidates. The small-R jets, b_1 and b_2, are chosen such that the absolute value of the mass difference between the hadronic and leptonic VLQ candidates, $|\Delta m| = |m_T^{\text{lep}} - m_T^{\text{had}}|$, is a minimum.

This reconstruction technique is optimized for BR($T \to Wb$) = 1 using simulated $T\bar{T}$ and $t\bar{t}$ events. The small-R jets paired with the W-candidates are referred to as the b-

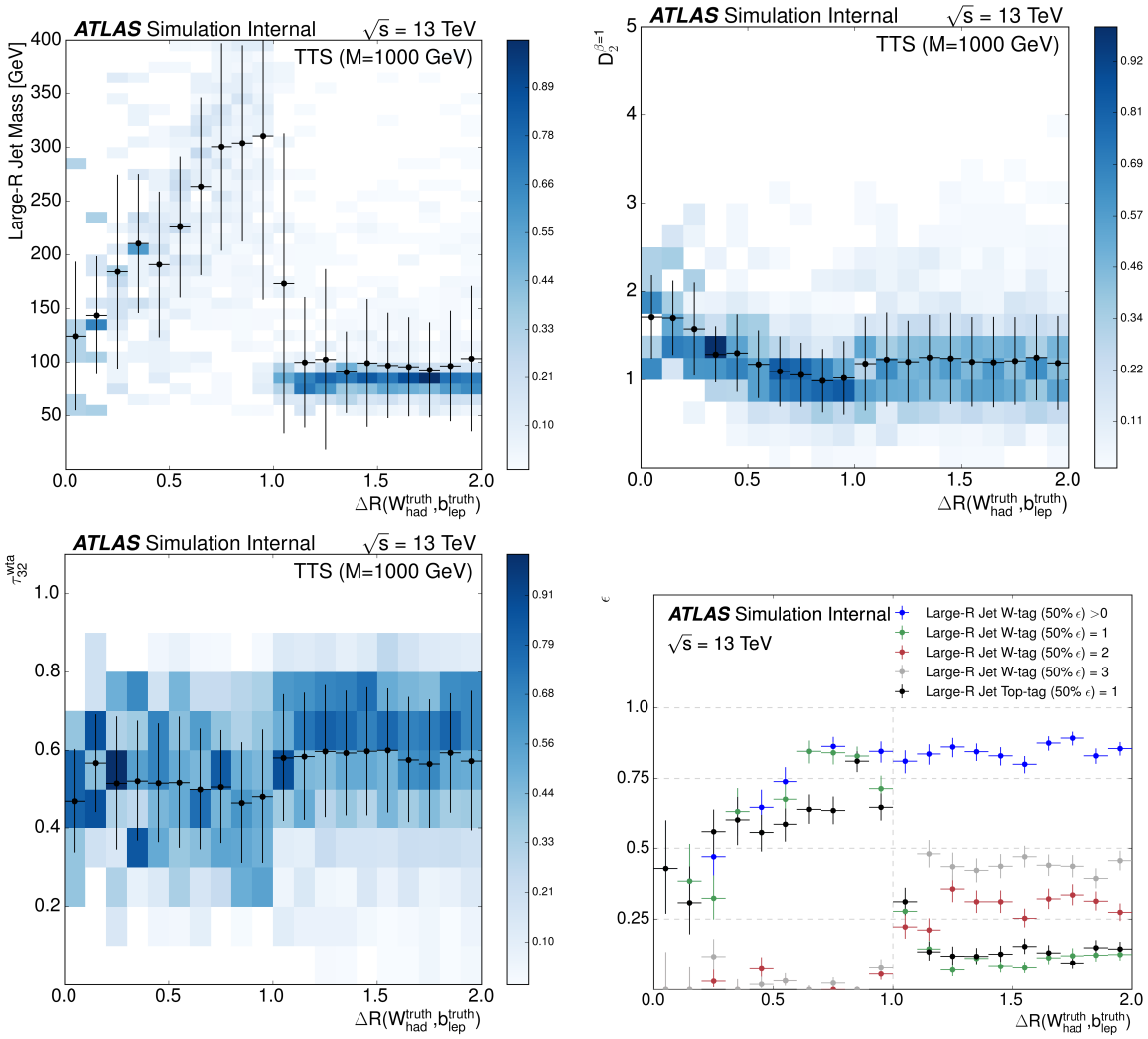

Figure 5.3: Mass (top left) and substructure $D_2^{\beta=1}$ (top right) and τ_{32}^{wta} (bottom left) distributions in the truth-matched large-R jet. The columns of each histogram are normalized and the label 'TTS' refers to singlet $T\bar{T}$ samples. The W– and top–tagging efficiencies are shown in the bottom right. Each distribution is presented as a function of the separation between the truth b-quark and hadronically-decaying W boson in $T\bar{T}$ ($m_T = 1$ TeV; $\text{BR}(T \to Wb) = 1$) events. The black points and error bars in the substructure distributions represent the mean and RMS, respectively, in each bin of $\Delta R(W_{\text{had}}^{\text{truth}}, b_{\text{lep}}^{\text{truth}})$.

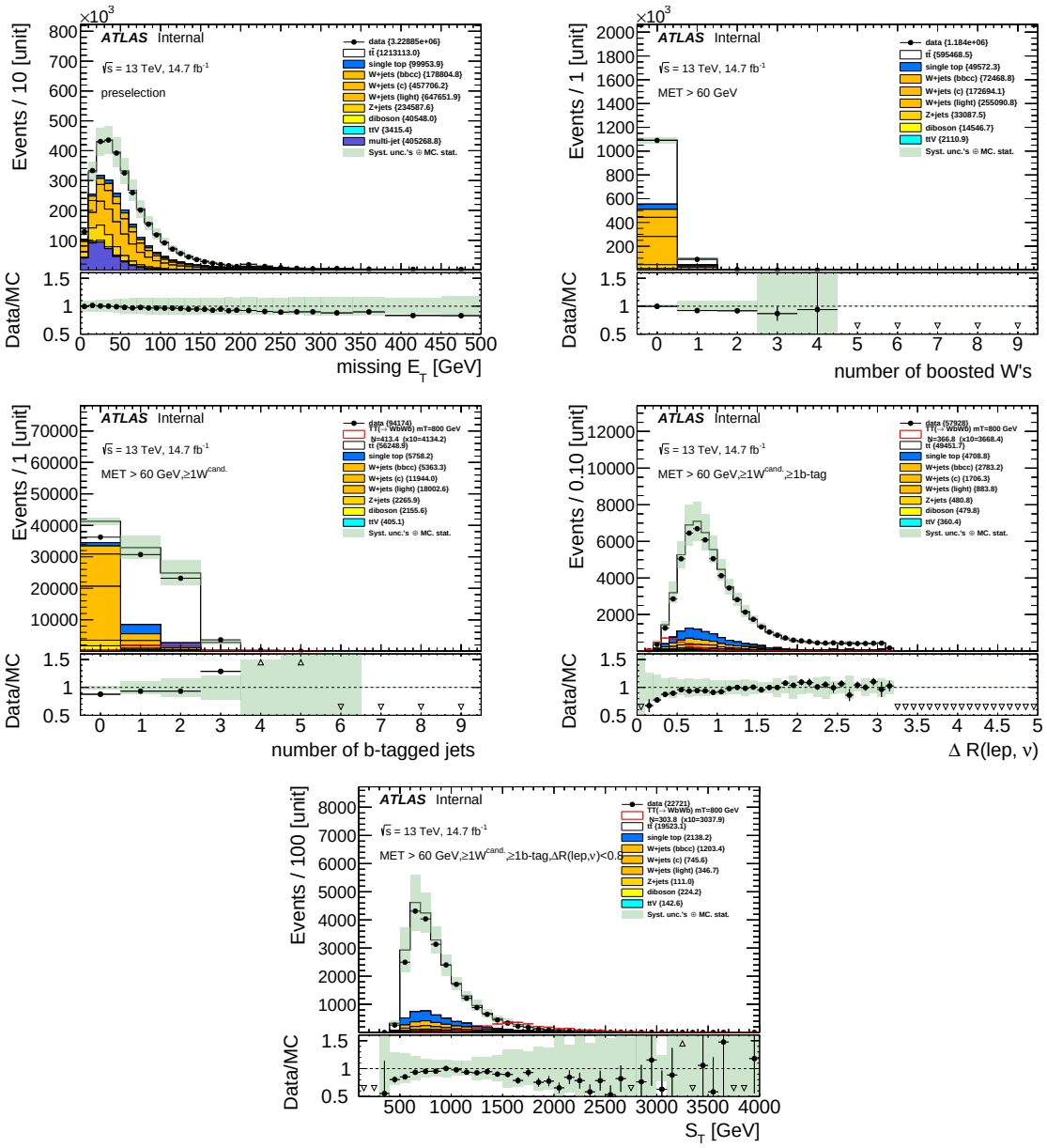

Figure 5.4: The figures depict all kinematic cut variables that define the boosted signal region. Each variable is shown just before the cut is applied, incorporating all previous cuts, as detailed in the text. The transverse momentum, the number of large-R jets selected as boosted hadronic W's, the number of b-tagged jets, $\Delta R(\ell, \nu)$ and S_{T} are illustrated. The green uncertainty band includes MC statistics as well as the $t\bar{t}$ modeling uncertainties.

quark candidates. If the event contains only 1 b-tag, the b-tagged jet is always considered as a b-quark candidate and permutations through the remaining small-R jets determines the the other b-quark candidate and the respective assignment of each b-quark candidate to the W-candidates. If the event contains 2 or more b-tags, the two leading p_{T} b-tagged jets are assumed to be both of the b-quark candidates and $|\Delta m|$ is used to assign each b-quark

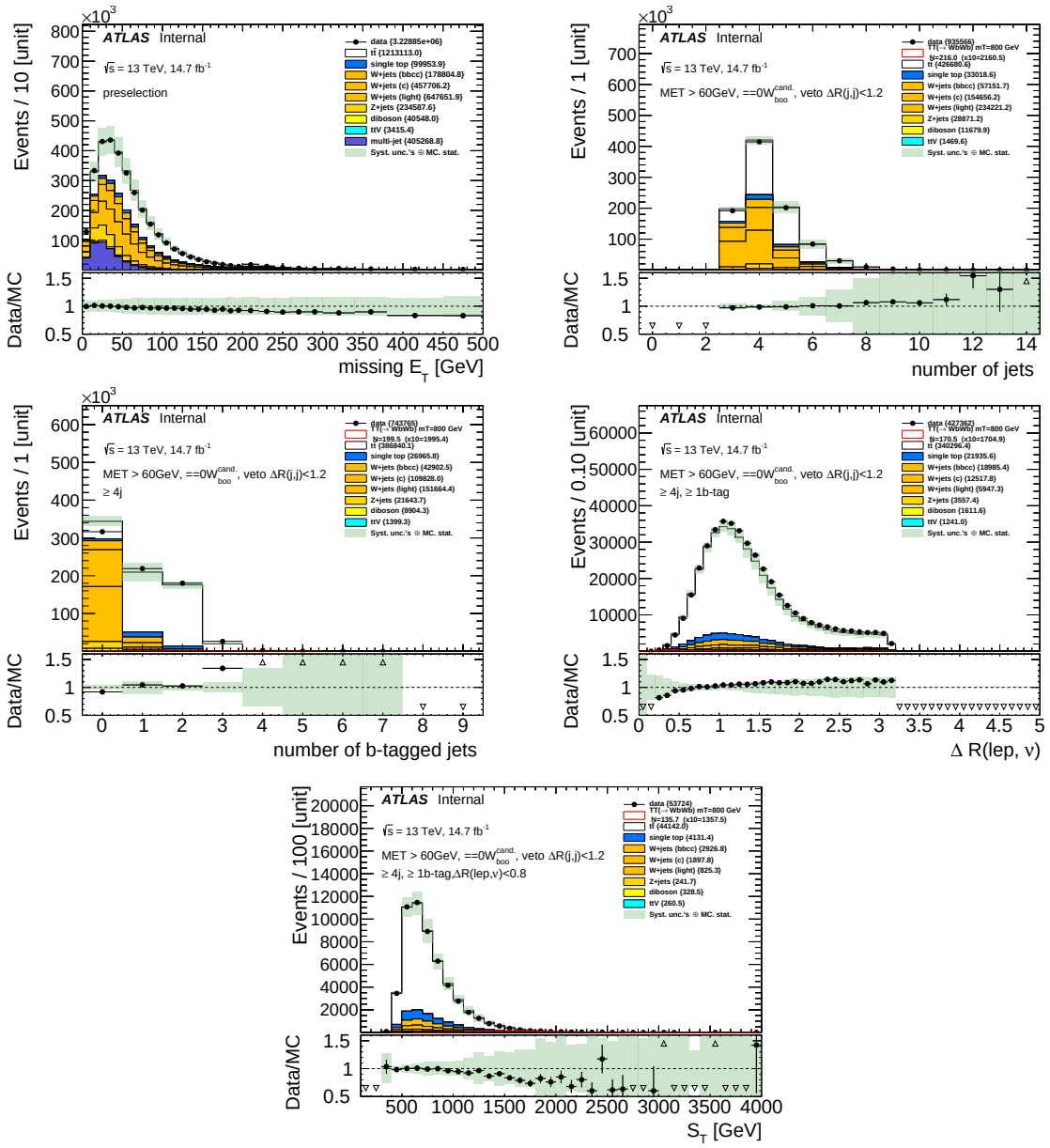

Figure 5.5: The figures depict all kinematic cut variables that define the resolved signal region. Each variable is shown just before the cut is applied, incorporating all previous cuts, as detailed in the text. The transverse momentum, the number of boosted hadronic W candidates, the number of b-tagged jets, $\Delta R(\ell, \nu)$ and S_{T} are illustrated. The green uncertainty band includes MC statistics as well as the $t\bar{t}$ modeling uncertainties.

candidate to one of the W-candidates. In Fig. 5.7 the reconstructed leptonic VLQ mass is shown for the boosted and resolved signal regions.

In simulated $T\bar{T}$ ($m_T = 1$ TeV), 45% of events have either of the b-quark candidates, b_i, located inside the hadronic W candidate, $\Delta R(b_i, W_{\mathrm{had}}) < 1.0$. As a performance metric, this algorithm is compared with the run 1 algorithm in Fig. 5.6. In the run 1 analysis, the

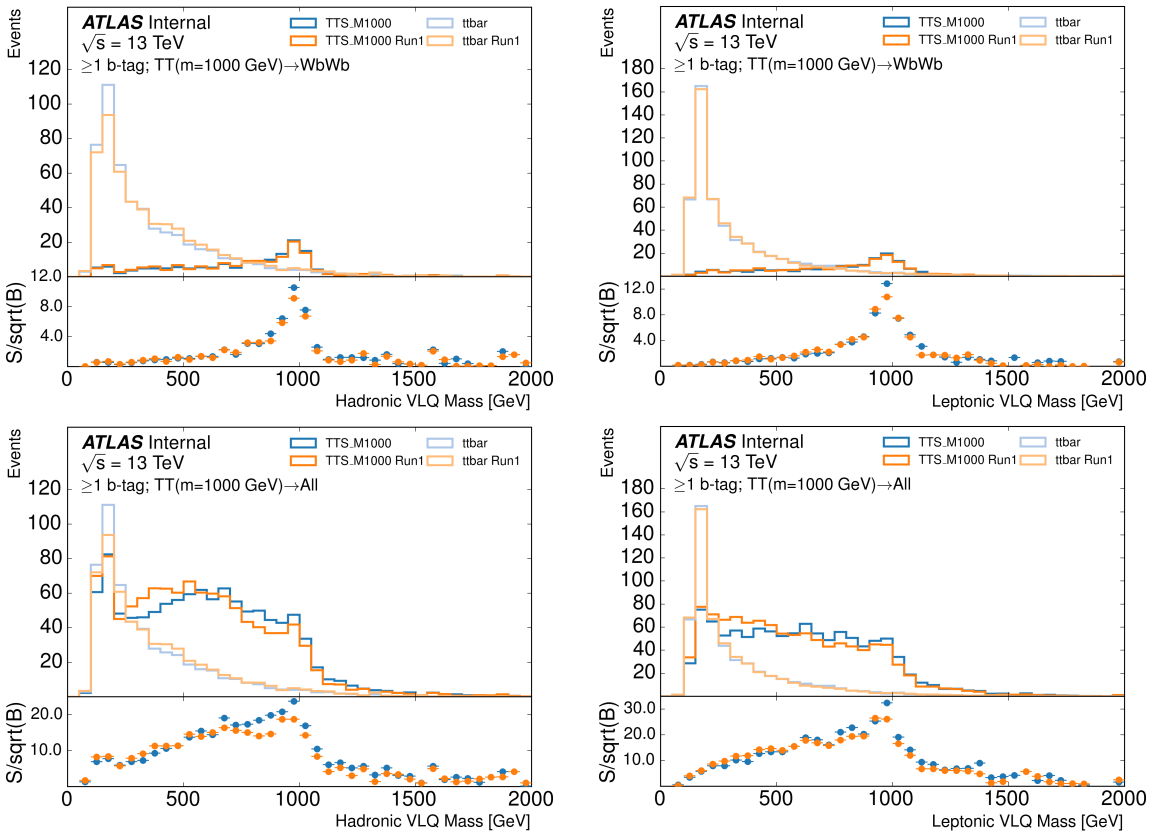

Figure 5.6: Reconstructed VLQ masses comparing the current run-2 method with the run-1 method. The bottom two figures show the mass reconstruction when considering all VLQ final state decays.

b-quark candidates were chosen as the two jets with the highest b-tagging discriminant [97][2]. The current method noticeably outperforms the approach from run 1 as shown in Fig. 5.6 for both m_T^{had} and m_T^{lep}. The subplot of Fig. 5.6 displays the significance, $S/\sqrt{B}$, in each bin as a quantitative measure of performance to compare the different methods. The significance at m_T^{lep}= 1 TeV is approximately 10 times greater than m_T^{lep}= 500 GeV for $T\bar{T} \to WbWb$. The significance of the inclusive final state decays of $T\bar{T}$ only improves by a factor of approximately 2 from m_T^{lep}= 500 GeV to m_T^{lep}= 1 TeV.

5.1.3 Classification of Event Topologies

Two orthogonal regions are defined to investigate the modeling of the $t\bar{t}$ background and understand its impact in the boosted and resolved signal regions, defined in Sect. 5.1.2. The

[2]For these comparisons, the 'run-1' method uses the `mv2c10` discriminant rather than the run 1 discriminant because it is neither recommended nor currently available for analyses.

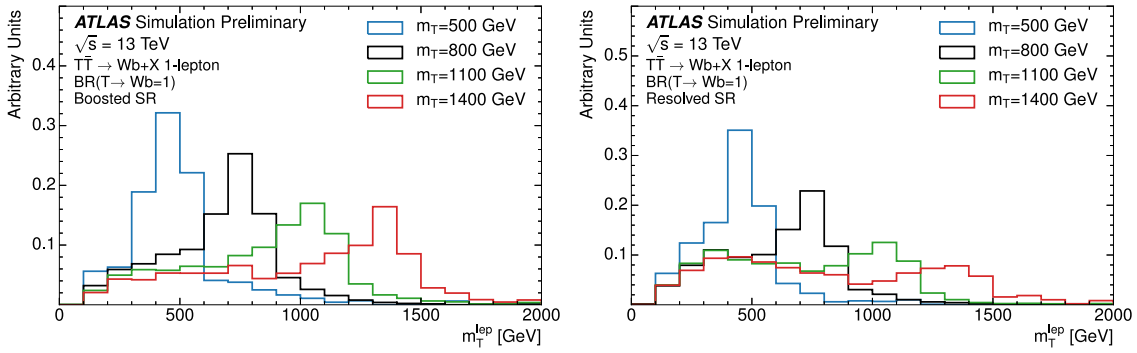

Figure 5.7: The reconstructed leptonic VLQ mass is shown for mass points of 500, 800, 1100, and 1400 GeV in the boosted (left) and resolved (right) signal regions. Both plots are normalized to unity for comparison of the relative shapes at each mass point.

two $t\bar{t}$ control regions were chosen to border the boosted signal region in an appropriate kinematic phase space, the $S_{\mathrm{T}} - \Delta R(\ell, \nu)$ plane.

The first of the $t\bar{t}$ control regions, referred to as CR1, applies an identical event selection as the final boosted signal region selection with two modifications:

- $\Delta R(\ell, \nu) > 1.0$
- $S_{\mathrm{T}} > 700$ GeV

This region is orthogonal to the signal region in $\Delta R(\ell, \nu)$ with a lower S_{T} threshold to maintain sufficient $t\bar{t}$ statistics. Fig. 5.8 shows plots of various kinematic distributions in the $t\bar{t}$ CR1.

The second of the $t\bar{t}$ control regions, referred to a CR2, also modifies the boosted event selection in the S_{T}vs. $\Delta R(\ell, \nu)$ plane:

- $\Delta R(\ell, \nu) < 1.0$
- $750 < S_{\mathrm{T}} < 1200$ GeV

This region is chosen to be as close to the boosted signal region as possible in terms of ΔR, while maintaining orthogonality with the signal region in S_{T}. The lower boundary on the S_{T} selection is set to be just above the region overwhelmingly dominated by $t\bar{t}$, as shown in Fig. 5.10. Fig. 5.9 shows plots of various kinematic distributions in the $t\bar{t}$ CR2.

To visualize the relationship between the control and signal regions, Fig. 5.10 shows two dimensional histograms in the S_{T} vs. $\Delta R(\ell, \nu)$ plane for $T\bar{T}$ ($m_T = 800$ GeV); BR($T \to Wb$) = 1 and the SM $t\bar{t}$ background. Table 5.3 lists the yields in data and MC for the three $t\bar{t}$ regions of interest.

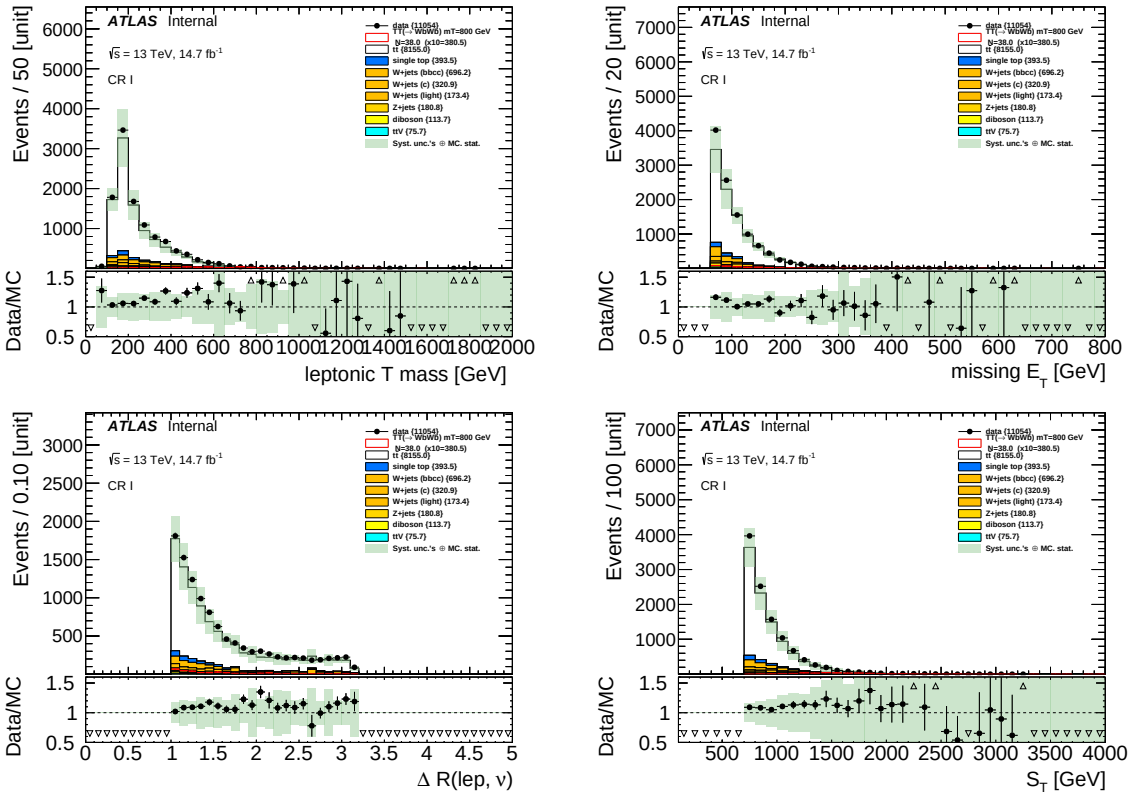

Figure 5.8: Data and MC comparison in CR1. The green band represents MC statistical uncertainties, as well as all $t\bar{t}$ modeling and detector-related systematic uncertainties.

The signal acceptance (fraction of total events that pass the full selection) is shown in Fig. 5.11 for the $SU(2)$ singlet and BR($T \to Wb$) = 1 scenario. Each bin corresponds to a single mass point; from 700-1200 GeV there are mass points separated by 50 GeV, with additional mass points at 500, 600, 1300, and 1400 GeV. The acceptances are drawn independently for each signal and control region and with the total acceptance for comparison.

For both the $SU(2)$ singlet and the BR($T \to Wb$) = 1 scenarios, the boosted signal region achieves the largest signal acceptance, followed by the resolved signal region. The $t\bar{t}$ CR2 achieves the lowest signal acceptance due to the lower S_{T} threshold, disfavored by high mass VLQs, while the $t\bar{t}$ CR1 achieves moderate signal efficiency because it complements the signal region with high S_{T} but a resolved leptonic W selection. Because the analysis has been optimized assuming BR($T \to Wb$) = 1, the $SU(2)$ singlet scenario has the lowest acceptance.

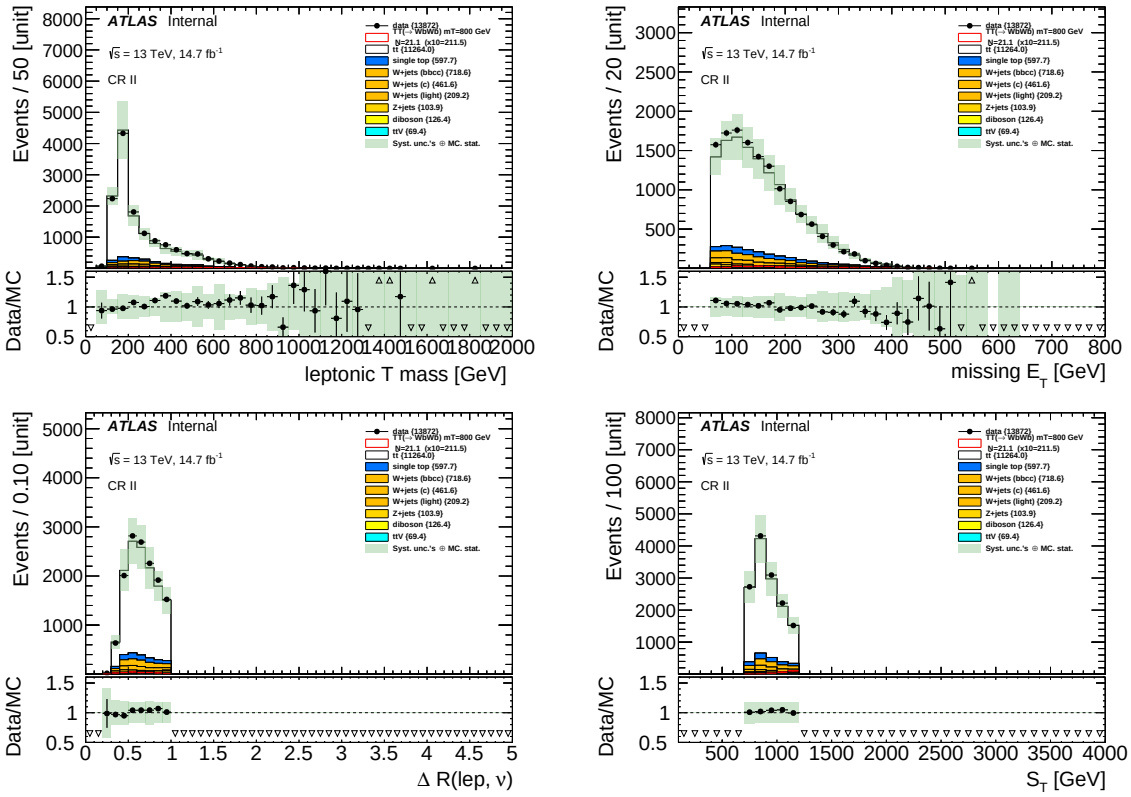

Figure 5.9: Data and MC comparison in CR2. The green band represents MC statistical uncertainties, as well as all $t\bar{t}$ modeling and detector-related systematic uncertainties.

5.1.4 Background Estimation

SM processes including single top, $t\bar{t}+V$, diboson, and Z+jets are estimated using simulated events normalized to the data luminosity and theoretical cross sections as described in Sect. 5.1.1. The estimation of the dominant $t\bar{t}$ background uses simulated events in dedicated control regions, discussed in Sect. 5.1.3. The W+jets background normalization and uncertainties are based on the flavor composition of the jets produced alongside the W boson and the multi-jet background is derived using a data-driven technique.

5.1.4.1 W+jets Normalization

Events with a leptonically decaying W boson in association with jets contributes the second largest background in the signal region. The theory cross section for this background has been computed to NNLO in QCD [98]. The charge asymmetry method exploits the difference between production of W^+ and W^- bosons in pp collisions. W boson production through $u\bar{d} \to W^+$ and $d\bar{u} \to W^-$ relies on the parton distribution function $u(x_1)\bar{d}(x_2)$ and $d(x_1)\bar{u}(x_2)$, respectively. These two sets of functions are different for protons, thus leading

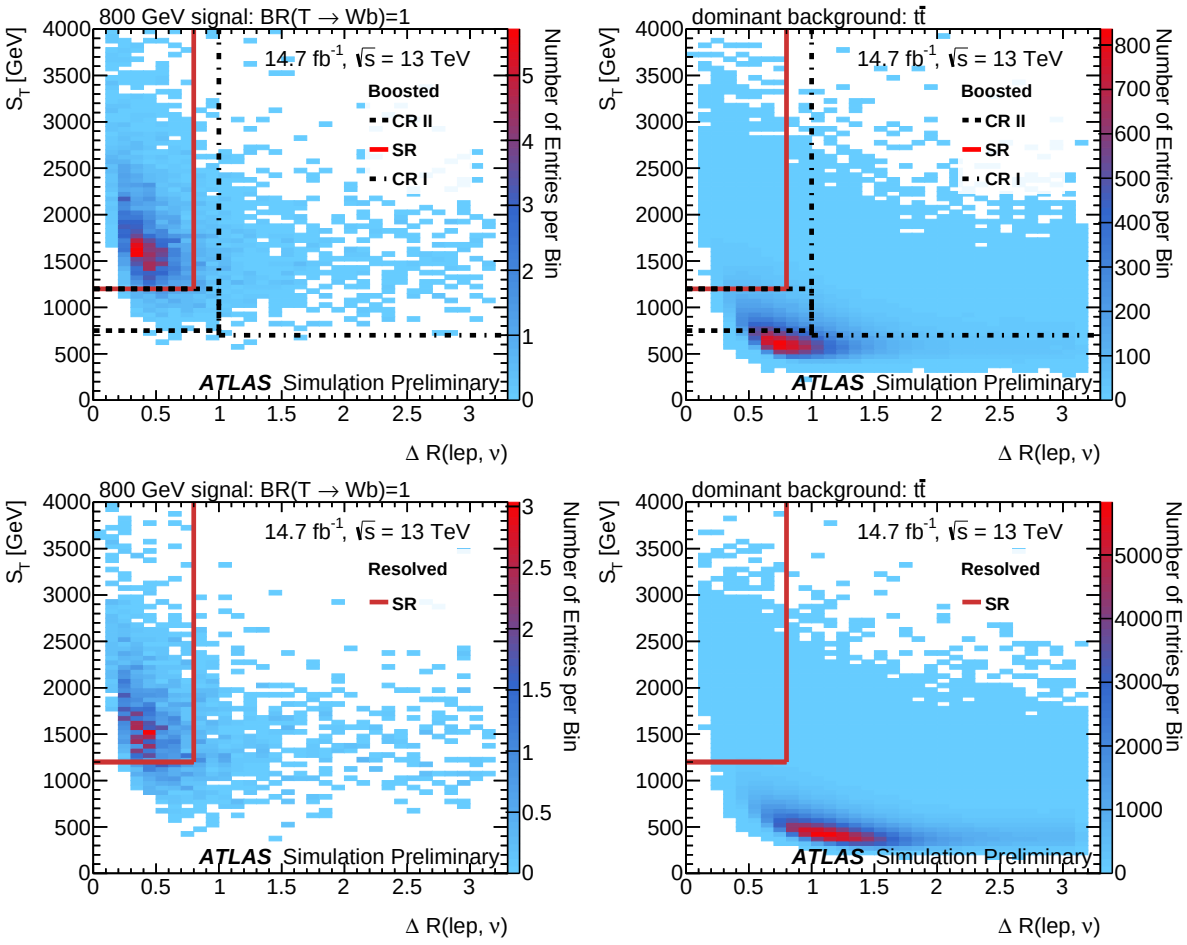

Figure 5.10: Two-dimensional histograms comparing the $T\bar{T}$ ($m_T = 800$ GeV) for BR($T \to Wb$) = 1 to SM $t\bar{t}$ in the boosted (top) and the resolved (bottom) regions in the S_{T} vs. $\Delta R(\ell, \nu)$ plane.

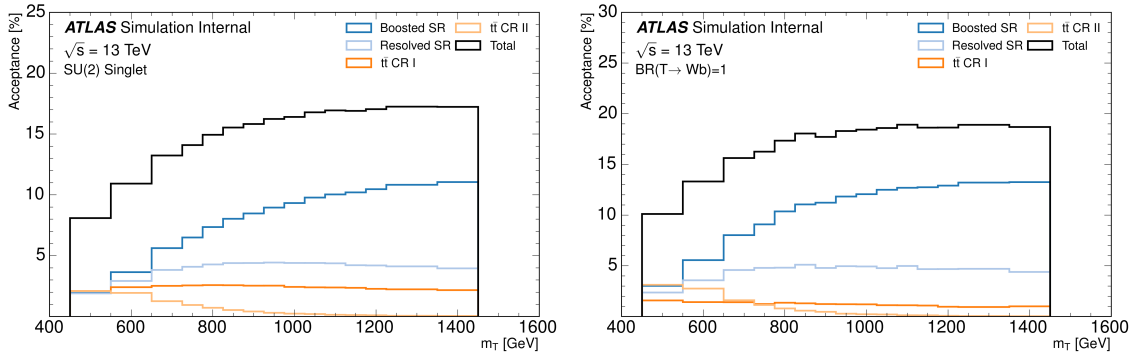

Figure 5.11: The signal acceptance at each VLQ mass for the signal and control regions. Individual selections are drawn separately to compare relative contributions to the total signal acceptance. Each bin represents a different generated VLQ mass.

Table 5.3: Observed and expected number of events in the $t\bar{t}$ control regions CR1 and CR2.

Sample	Yield (CR1)		Yield (CR2)	
$t\bar{t}$	8200	±2700	11 300	±3200
Single top	390	±40	600	±70
W+jets	1190	±220	1390	±200
Z+jets	180	±50	100	±50
Diboson	110	±130	130	±170
$t\bar{t}+V$	76	±7	69	±7
Multi-jet	450	±450	360	±360
Total	10 600	±2600	13 900	±2500
TTS (m_T=800 GeV)	38	±4	21	±3
Data	11 054		13 872	

to an asymmetry in the production of W^+ and W^- bosons at the LHC. W+jets contribution such that the measured asymmetry between positive and negative lepton events agrees with the prediction. The total number of W+jets events in data, $N_W = N_{W^+} + N_{W^-}$, can be estimated based on the difference between positively-charged and negatively-charged W bosons, $N_{W^+} - N_{W^-}$, and their ratio, r_{MC}, determined from simulation, according to:

$$N_W = \left[\frac{r_{MC}+1}{r_{MC}-1}\right](N_{W^+} - N_{W^-})_{\mathrm{measured}} \tag{5.5}$$

This method does not allow for different asymmetries or different scale factors for W+light jets versus $W + c/c\bar{c}$ jets versus $W + b/b\bar{b}$ events. For this analysis, it is most important that the heavy flavor component ($W + c/c\bar{c}$ and $W + b/b\bar{b}$) events are scaled properly, which could be significantly different from what the scaling needed for the W+light jets component. To determine separate scaling factors for heavy flavor and light jet components, the charge asymmetry requirement is applied to events with zero b-tags or at least one b-tag. This leads to two equations with two unknowns, the scale factor for W+light jets (K_{light}) and the scale factor for W+heavy flavor events (K_{hf}) that can be analytically determined Eq. 5.6.

$$\begin{aligned} K_{\mathrm{light}} &= \frac{(\Delta d_0 - \Delta b_0)\Delta h_1 - (\Delta d_1 - \Delta b_1)\Delta h_0}{\Delta l_0 \Delta h_1 - \Delta l_1 \Delta h_0} \\ K_{\mathrm{hf}} &= \frac{(\Delta d_0 - \Delta b_0)\Delta l_1 - (\Delta d_1 - \Delta b_1)\Delta l_0}{\Delta h_0 \Delta l_1 - \Delta h_1 \Delta l_0} \end{aligned} \tag{5.6}$$

where Δd, Δb, Δl, and Δh are the difference between the number of e^+ and e^- events for data, non-W asymmetric backgrounds, W+light jet events, and W+heavy flavor jet events, respectively, and the subscript refers to the number of b-tags.

The scale factors were estimated for two selections: the baseline and the baseline + at least one large-R jet with $p_{\mathrm{T}} > 200$ GeV. Both selections are kinematically and topologically different from the signal regions, thus the scale factors are not applied in this analysis, which are summarized in Table 5.4. Applying tighter selection results in unstable estimates for scale factors and uncertainties due to limited statistics. To account for this, the heavy flavor component, which dominates in the signal region, receives a conservative uncertainty of 50%, to cover differences listed in Table 5.4.

Table 5.4: Scale factors for W+light jet and W+heavy flavor events derived with different selections.

Selection	K_{light}	K_{hf}
Baseline	0.70 ± 0.09	0.94 ± 0.17
Baseline + ≥ 1 large-R jet ($p_{\mathrm{T}} > 200$ GeV)	0.78 ± 0.14	1.08 ± 0.28

5.1.4.2 Multi-jet Estimation

The QCD multi-jet contribution to the selected events occurs when a a jet or a photon is mis-identified as an electron or from the selection of a non-prompt lepton, e.g., a semi-leptonic $b-$ or $c-$hadron decay. The multi-jet background is expected to be small in the signal region due to the tight requirements on S_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$.

A data-driven 'matrix method' (MM) [99, 100] technique is used to estimate this background. The MM exploits differences in lepton identification-related properties between prompt, isolated leptons from W and Z boson decays ("real" leptons) and those where the lepton candidates are either non-isolated or result from the mis-identification of photons or jets ("fake" leptons). Two samples are defined which differ only in the lepton identification criteria: a "tight" sample and a "loose" sample, the former a subset of the latter. The requirements for loose and tight leptons are given in Table 5.5.

Table 5.5: Loose and tight lepton identification requirements for the multi-jet estimation. See Sect. 4.2 for details on the lepton identification.

	Tight Electrons	Loose Electrons	Tight Muons	Loose Muons
ID Quality:	Tight likelihood	Medium likelihood	Medium	Medium
Isolation:	Gradient	None	Gradient	None

The MM assumes that the number of loose (tight) selected events in each sample N^{Loose} (N^{Tight}) can be expressed as a linear combination of the numbers of events with real and fake leptons, such that

$$\begin{aligned} N^{\text{Loose}} &= N_{\text{r}}^{\text{Loose}} + N_{\text{f}}^{\text{Loose}} \\ N^{\text{Tight}} &= \epsilon_{\text{r}} N_{\text{r}}^{\text{Loose}} + \epsilon_{\text{f}} N_{\text{f}}^{\text{Loose}} \end{aligned} \tag{5.7}$$

where $\epsilon_{\text{r(f)}}$ is the efficiency for a real (fake) lepton to satisfy the tight criteria if it already satisfies the loose criteria. The efficiencies are measured for one and two-dimensional parameterizations in lepton p_{T} (p_{T}^{ℓ}) and η^{ℓ}, $E_{\text{T}}^{\text{miss}}$, separation between the lepton and jets ($\Delta\phi$), and b-tag multiplicity (N_b). Products of the one- and two-dimensional parameterizations are used to construct the final efficiencies used for calculating the MM event weights. Table 5.6 shows the efficiency calculations for each lepton flavor. Plots for baseline selection with the multi-jet estimate from the MM are shown in Fig. 5.12.

Table 5.6: Parametrization used for real and fake lepton efficiencies used to calculate the MM weights. A smooth transition between the low and high-p_{T} efficiencies is achieved using a sigmoid function with a centered at 150 GeV with a width of 20 GeV.

	Fake Efficiency	Real Efficiency
Electrons:	$\sqrt[3]{\epsilon_{\text{f}}(p_{\text{T}}^{\ell}, N_b) \cdot \epsilon_{\text{f}}(\Delta\phi, \eta^{\ell}) \cdot \epsilon_{\text{f}}(p_{\text{T}}^{\ell}, \Delta\phi)}$	$\epsilon_{\text{r}}(p_{\text{T}}^{\ell})$
Muons (low-p_{T}):	$\sqrt{\epsilon_{\text{f}}(p_{\text{T}}^{\ell}, \Delta\phi) \cdot \epsilon_{\text{f}}(E_{\text{T}}^{\text{miss}})}$	$\sqrt{\epsilon_{\text{r}}(p_{\text{T}}^{\ell}) \cdot \epsilon_{\text{r}}(\Delta\phi)}$
Muons (high-p_{T}):	$\epsilon_{\text{f}}(p_{\text{T}}^{\ell})$	$\epsilon_{\text{r}}(p_{\text{T}}^{\ell})$

Eq. 5.7 can be re-written to express the number of tight events coming from fake leptons:

$$N_{\text{f}}^{\text{Tight}} = \frac{\epsilon_{\text{f}}}{\epsilon_{\text{r}} - \epsilon_{\text{f}}}(\epsilon_{\text{r}} N^{\text{Loose}} - N^{\text{Tight}}) \tag{5.8}$$

which can be generalized to define a weight w_i for each event, i, in data

$$w_i = \frac{\epsilon_{\text{f}}}{\epsilon_{\text{r}} - \epsilon_{\text{f}}}(\epsilon_{\text{r}} - \delta_i) \tag{5.9}$$

where δ_i is 1 if the event passes the tight selection and 0 otherwise. With Eq. 5.9, any histogram can be defined for the multi-jet background using properties of the events in data with the weight w_i. The MM performs best in regions with high statistics, e.g., the baseline selection. In the signal region, however, the multi-jet contribution is very low. In such

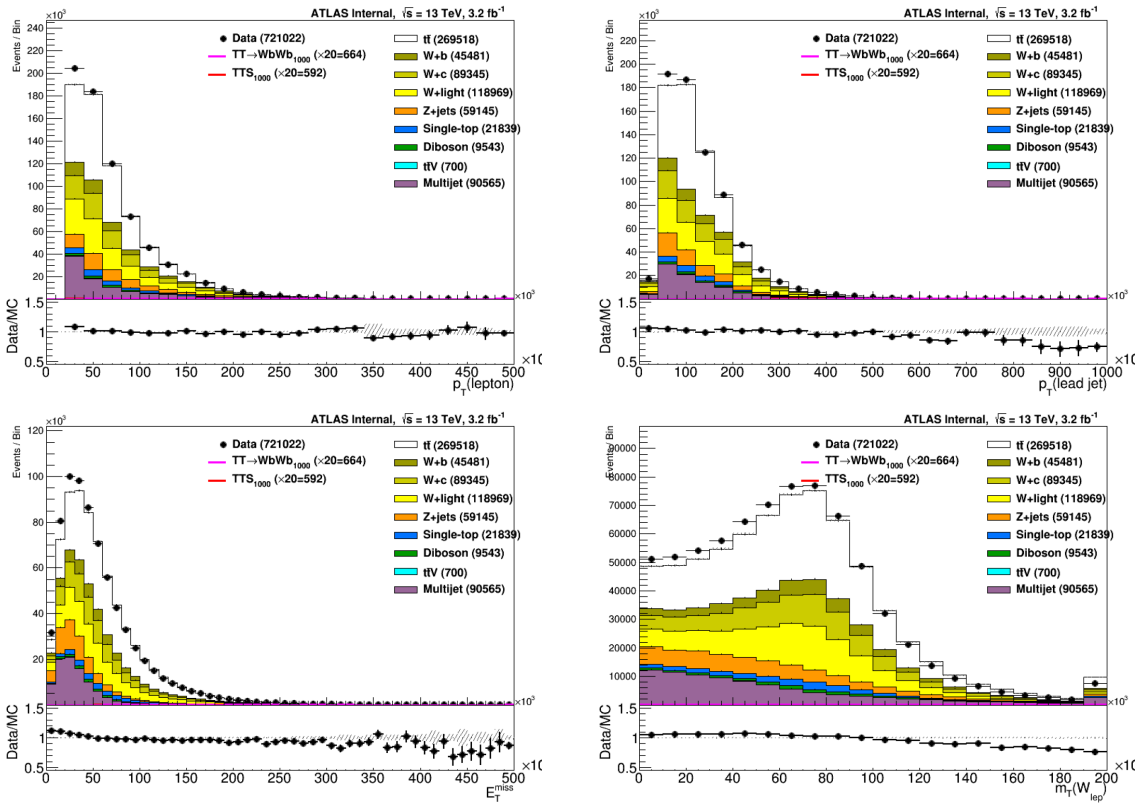

Figure 5.12: Plots of the lepton p_{T}, leading jet p_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, and transverse W mass distributions after the baseline selection showing the multi-jet estimate from the matrix method. No corrections have been applied to the W+jets or $t\bar{t}$ background normalizations. Only statistical uncertainties are shown.

a region of low statistics, the MM can give unstable distributions and bins with negative weights. For this reason, an ABCD method, described below, is used to cross-check the multi-jet estimation.

The ABCD method divides the analysis into a grid of four regions: a signal region D, a region with one cut inverted B, a region with a second cut inverted C, and a region with both cuts inverted A. The B, C, and A regions are used to estimate the normalization of the multi-jet background in region D, by the equation:

$$N_D = N_B \times \frac{N_C}{N_A} \tag{5.10}$$

where N_i (i = A, B, C, D) refers to the number of events in each region. The estimate is performed with this structure as well as with higher dimensionality, inverting different combinations of S_{T}, number of b-tags, $E_{\mathrm{T}}^{\mathrm{miss}}$, and $\Delta R(\ell, \nu)$. Across the different ABCD methods, 1.6–9.8 multi-jet events are predicted in the signal region compared to 19±9 events predicted by the MM. The statistics are too low for either method to provide a reliable shape

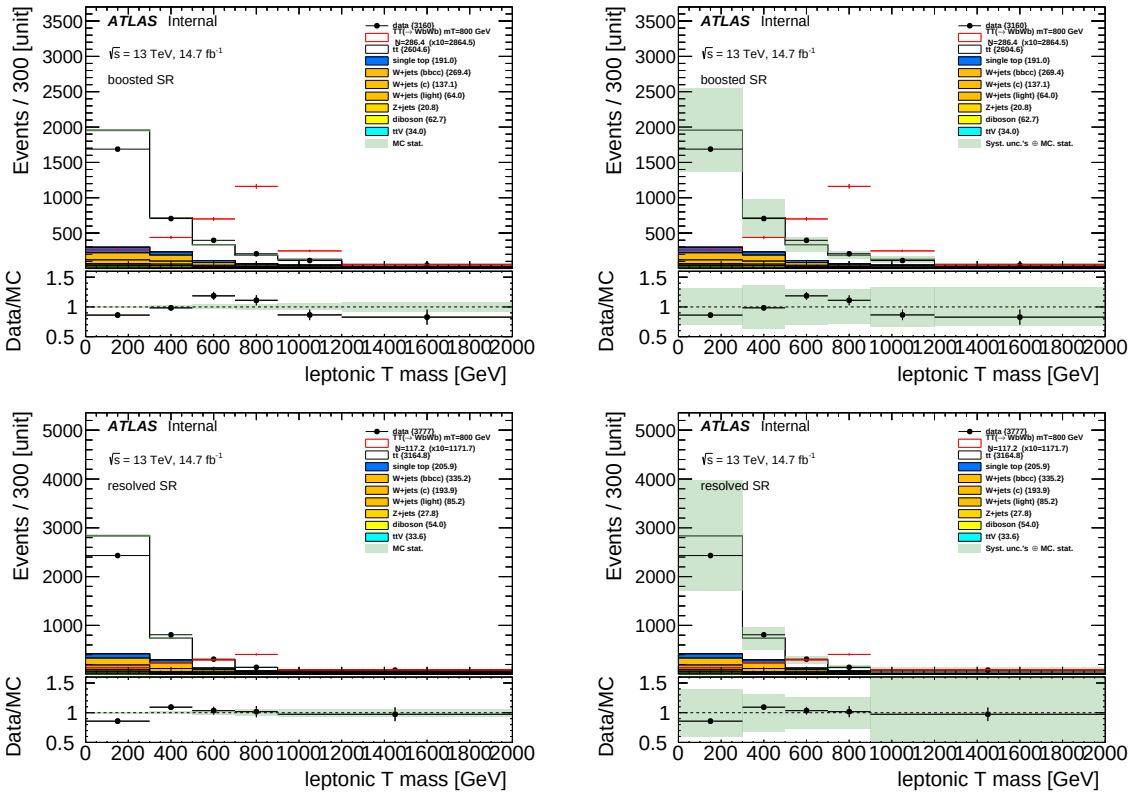

Figure 5.13: Comparison of data and MC for the boosted (top) and resolved (bottom) signal regions in the $\ell+$jets channel before (left) and after (right) the inclusion of the detector and $t\bar{t}$ background modeling systematics. No fake estimate is present in these comparisons.

in the signal region. Thus, the multi-jet normalization estimate in the signal region is 10±10 events, consistent with both estimation methods, and the multi-jet shape is set to the shape of the combined distribution of $t\bar{t}+V$, diboson, and Z+jets. A shape systematic uncertainty is defined by the difference between the shapes after the baseline selection.

5.1.5 Systematic Uncertainties

This section describes the sources of systematic uncertainty considered in this analysis. These uncertainties can be broken down into several broad categories: luminosity and cross section uncertainties, detector-related experimental uncertainties, uncertainties on data-driven background estimations, and modeling uncertainties on background processes. Each source of uncertainty is treated as a nuisance parameter in the fit of m_T^{lep}. The impact of the systematic uncertainties before the likelihood fit can be seen in Fig. 5.13, where m_T^{lep} is shown for the boosted and resolved signal regions with and without the systematic uncertainties described in this section.

5.1.5.1 Luminosity and Cross Section Uncertainties

The preliminary uncertainty in the combined 2015+2016 integrated luminosity is 2.9%. It is derived, following a methodology similar to that detailed in [101], from a preliminary calibration of the luminosity scale using $x - y$ beam-separation scans performed in August 2015 and May 2016.[3] This systematic uncertainty is applied to all backgrounds estimated using simulated samples and normalized to the measured integrated luminosity.

Theoretical cross section uncertainties have been used for the simulated samples that are used in this analysis. The uncertainties for W/Z+jets and diboson production are taken 5% and 6%, respectively [102, 103]. An additional 24% uncertainty is added in quadrature for each additional jet. Further, two uncertainties related to the HF content on W+jets background are considered and included as an additional shape uncertainty, keeping the overall W+jets template normalization fixed. The uncertainty is derived by individually varying the W+jets(b) and W+jets(c) components by 50%, see Sect. 5.1.4.1. For top pair and single top production the uncertainty is taken as 6% [44,104,105]. The data-driven multi-jet estimation, receives a conservative uncertainty of 100% is applied to the normalization as described in Sect. 5.1.4.2.

5.1.5.2 Detector-related uncertainties

Signal and backgrounds estimated using simulated events include systematic uncertainties on the detector simulation. These uncertainties are associated with the different physics objects: muons, electrons, $E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ and large-R jets, b-tagging, and pile-up, as described in Sect. 4.2.

The lepton systematics address uncertainties with the techniques used in the identification, trigger, and kinematics. The $E_{\mathrm{T}}^{\mathrm{miss}}$ systematic uncertainties address the uncertainty in the calculation of the track-based soft term scale and resolution. The jet uncertainties, both small-R and large-R, address the many components used in the calibration sequence. The small-R jet uncertainties are dedicated to JVT and energy scale and resolution [69]. The large-R jet uncertainties address uncertainties in the kinematics and substructure variables, $D_2^{\beta=1}$ and τ_{32}^{wta}, for a baseline uncertainty comparing the difference between data and prediction, comparison of different MC generators, considering uncertainties on the tracks, and the statistical uncertainty on the results. The substructure and kinematic uncertainties are assumed to be fully correlated.

[3] This paragraph is directly copied from the ATLAS luminosity for physics definition, as recommended by the ATLAS publications committee: `https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics`

The large-R jet resolution uncertainties are estimated at the analysis-level by smearing each variable of interest (p_{T}, mass, $D_2^{\beta=1}$, and τ_{32}^{wta}) and repeating the analysis chain. The p_{T} resolution uncertainty is estimated for each jet by resetting the jet p_{T} to a random value from a Gaussian distribution (μ = jet p_{T}, $\sigma = 0.02$) and propagating the result through the event selection. The resolution uncertainties for mass, $D_2^{\beta=1}$, and τ_{32}^{wta} are estimated using the jet response from simulated $t\bar{t}$ events. The large-R jet response for X is defined as $X_{\mathrm{reco}}/X_{\mathrm{truth}}$ where X_{reco} is the value of X for the reconstructed large-R jet and X_{truth} is the value of X for the truth large-R jet (also trimmed with the same parameters as the reconstructed large-R jet) geometrically matched to the reconstructed jet. The response is fit with a triple-Gaussian and the largest width is taken as the nominal width (σ_{nom}) of the response. For each jet, the parameter of interest is smeared with a Gaussian $G(X, \sigma')$ where X is the value of the mass, $D_2^{\beta=1}$, or τ_{32}^{wta} of the jet and $\sigma' = x \cdot \sigma_{\mathrm{nom}}$ where $x = 0.57$ for $D_2^{\beta=1}$ and τ_{32}^{wta} and $x = 0.66$ for mass. The values of x were derived to cover discrepancies observed in run 1 large-R jet distributions. The distributions in Figs. 5.14–5.17 compare the estimated resolution uncertainties with the statistical uncertainty in simulated $t\bar{t}$ events for the different analysis regions. In each region, the large-R jet resolution uncertainty is smaller than statistical uncertainty in nearly every bin of m_T^{lep} and thus it is considered negligible for this result.

The flavor-tagging terms describe uncertainties in the b-tagging efficiency, jet energy, $t\bar{t}$ modeling, and contamination from light and c-jets. In addition, an uncertainty is associated to the re-weighting procedure due to pile-up. The groups of uncertainties with the largest impact are those associated with the jet energy scale and flavor-tagging.

5.1.5.3 Top Quark Modeling Uncertainties

In the dominant $t\bar{t}$ background, systematic uncertainties on shape and normalization are taken into account for variations in extra radiation, choice of parton shower, and choice of matrix-element generator. All samples for assessing the systematic uncertainties are normalized to the same $t\bar{t}$ cross section. For each uncertainty, the shape of m_T^{lep} is examined.

The extra radiation uncertainty is estimated by comparing samples with more and less radiation to the nominal sample. Fig. 5.18 shows the impact on the nominal sample in simulated $t\bar{t}$ events in each control and signal regions.

The uncertainty due to the choice of matrix-element generator compares samples generated using `aMC@NLO` and POWHEG-BOX that both use PYTHIA 8 for parton showering. The samples compared are both generated using the full simulation, and the uncertainty is smoothed and symmetrized. This uncertainty is compared with the nominal sample is shown in Fig. 5.18 for simulated $t\bar{t}$ events in each control and signal regions.

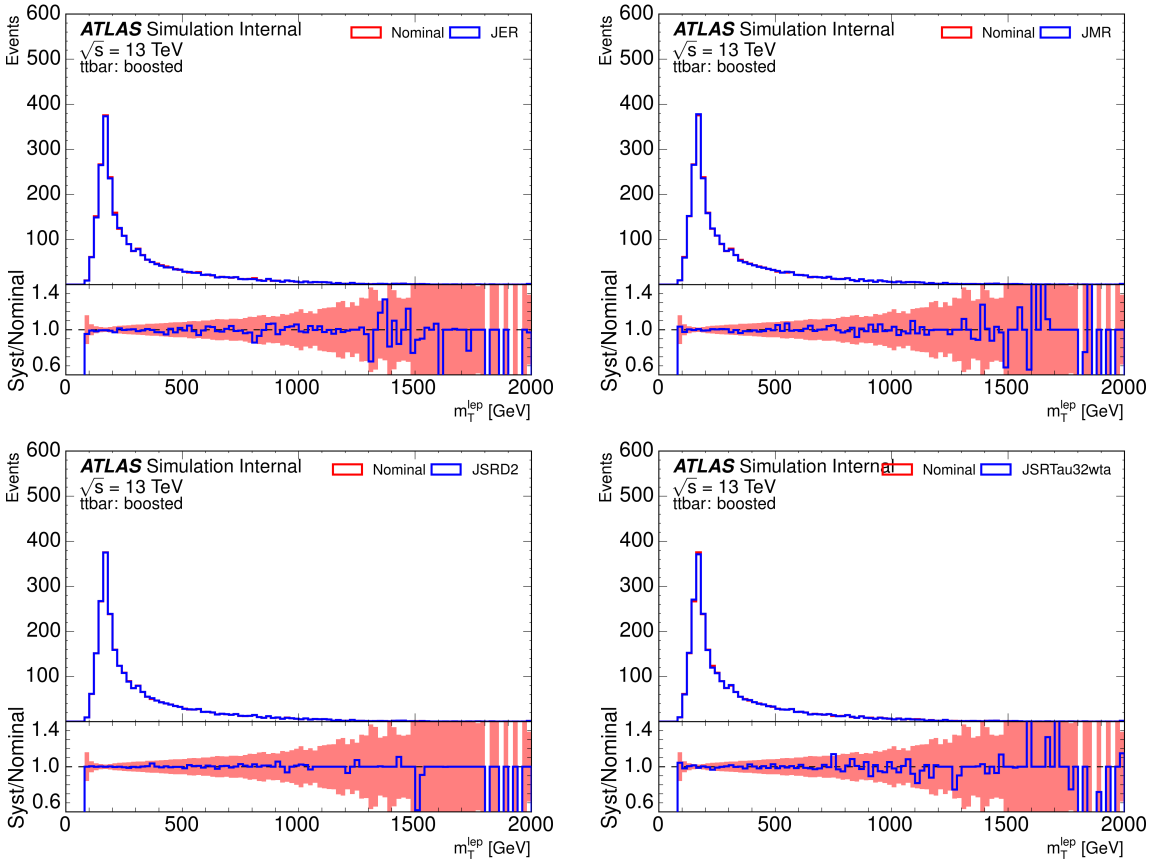

Figure 5.14: Distributions for the large-R jet nominal and resolution variations in the boosted signal region for m_T^{lep}. The red filled area represents the statistical uncertainty. The ratio between the systematic variation and nominal is taken as the resolution uncertainty for p_{T} (top left), mass (top right), $D_2^{\beta=1}$ (bottom left), and τ_{32}^{wta} (bottom right).

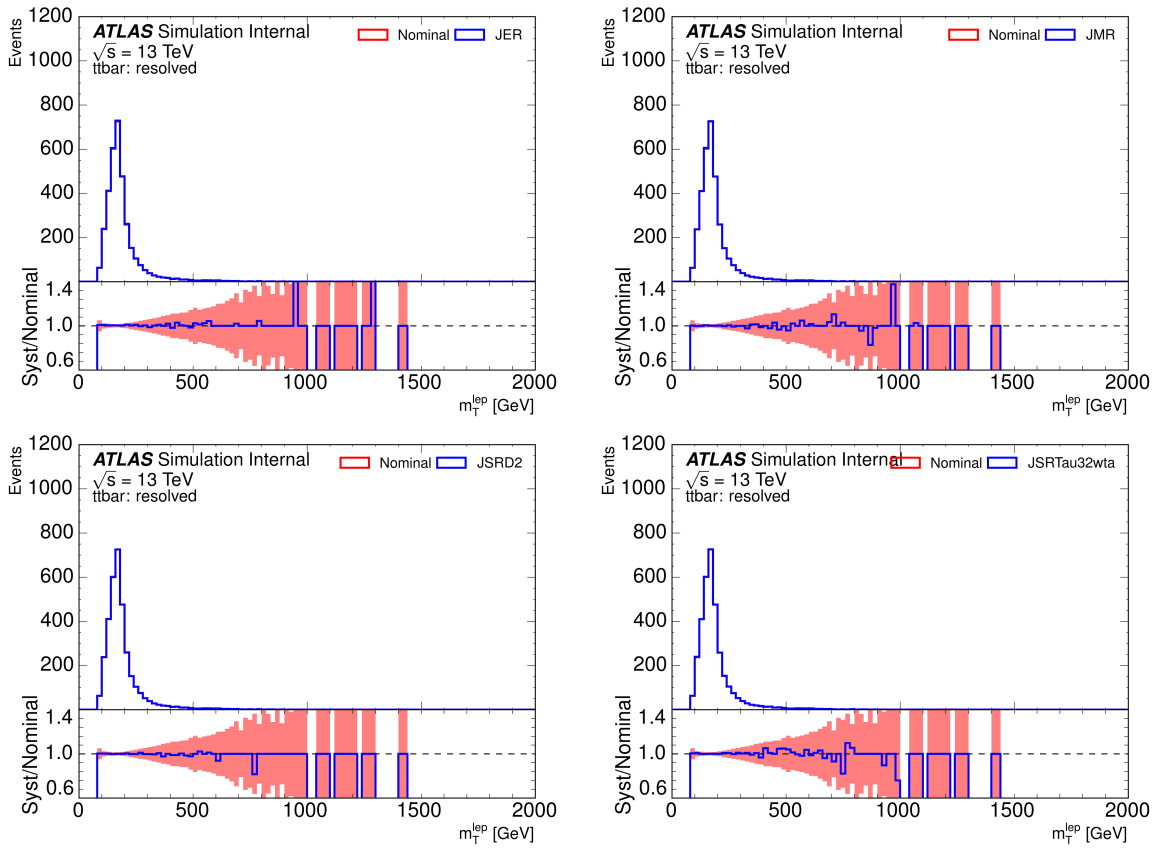

Figure 5.15: Distributions for the large-R jet nominal and resolution variations in the resolved signal region for m_T^{lep}. The red filled area represents the statistical uncertainty. The ratio between the systematic variation and nominal is taken as the resolution uncertainty for p_{T} (top left), mass (top right), $D_2^{\beta=1}$ (bottom left), and τ_{32}^{wta} (bottom right).

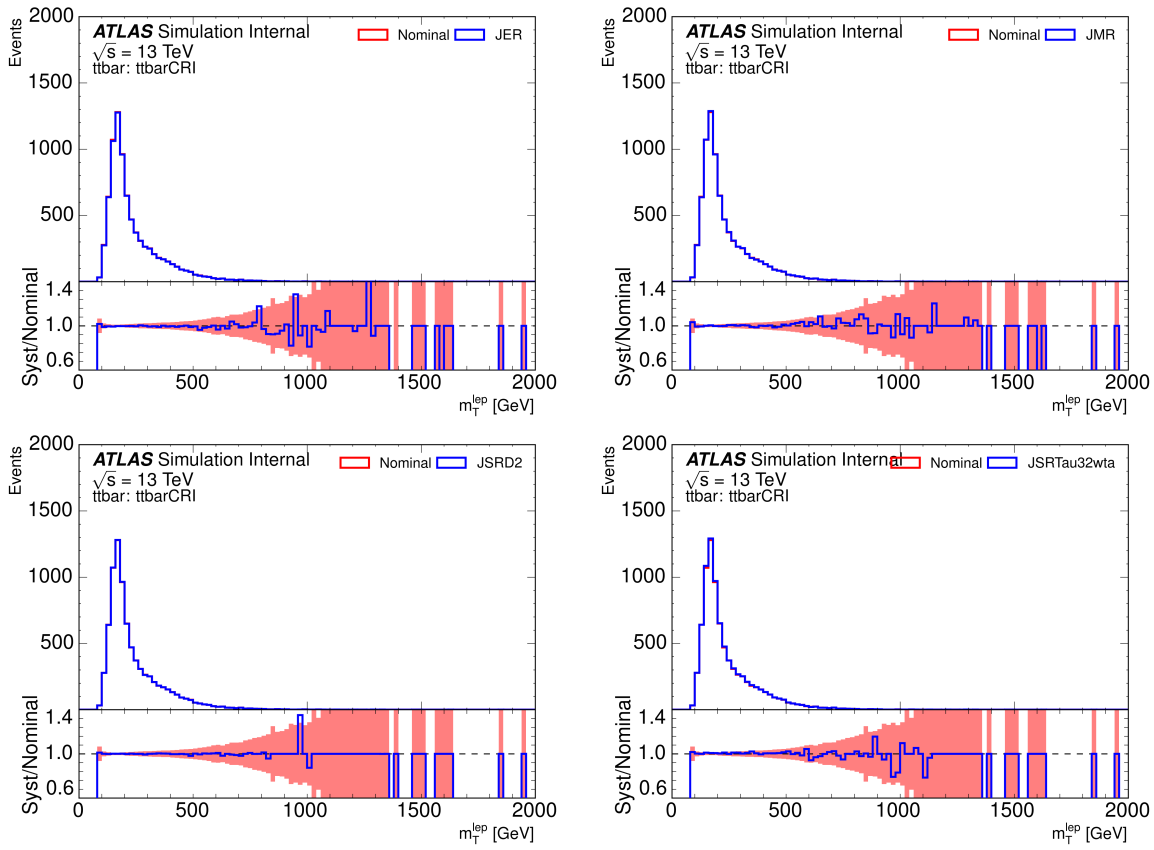

Figure 5.16: Distributions for the large-R jet nominal and resolution variations in CR1 for m_T^{lep}. The red filled area represents the statistical uncertainty. The ratio between the systematic variation and nominal is taken as the resolution uncertainty for p_{T} (top left), mass (top right), $D_2^{\beta=1}$ (bottom left), and τ_{32}^{wta} (bottom right).

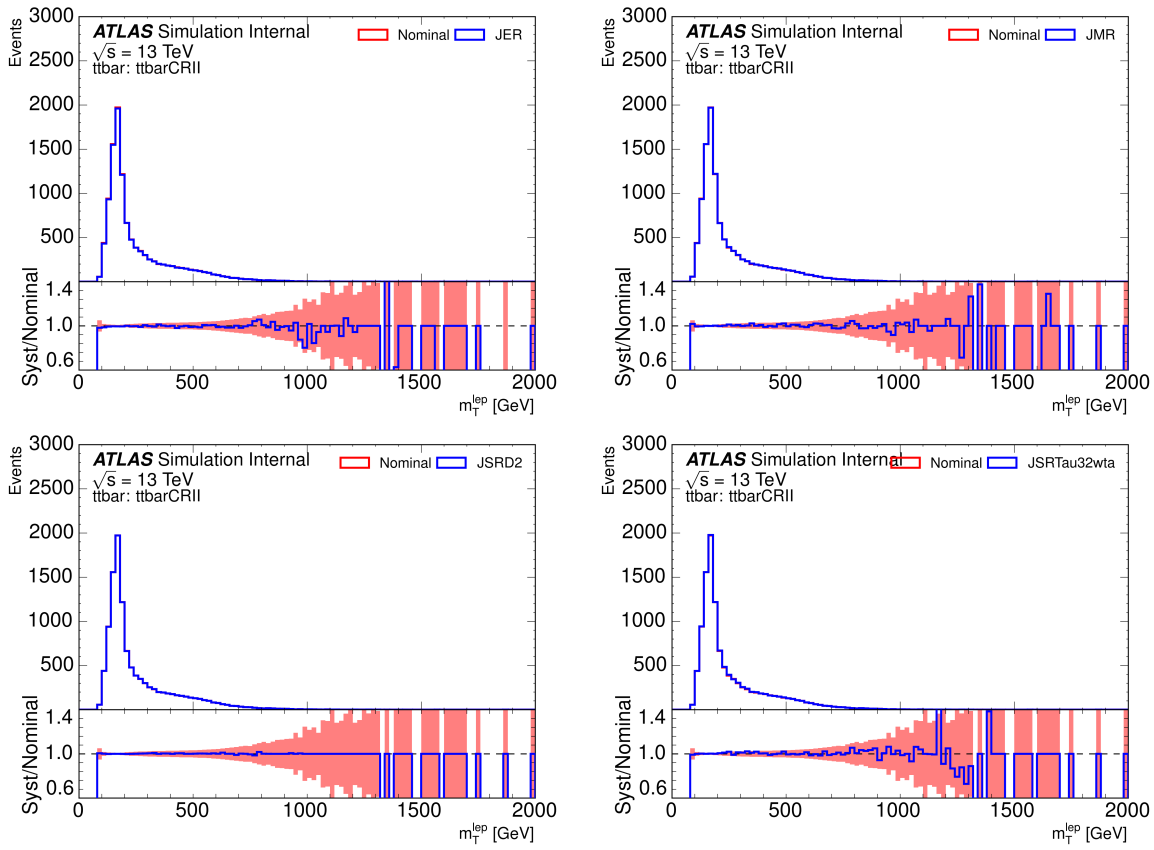

Figure 5.17: Distributions for the large-R jet nominal and resolution variations in CR2 for m_T^{lep}. The red filled area represents the statistical uncertainty. The ratio between the systematic variation and nominal is taken as the resolution uncertainty for p_{T} (top left), mass (top right), $D_2^{\beta=1}$ (bottom left), and τ_{32}^{wta} (bottom right).

The parton shower and underlying event uncertainties are evaluated by comparing the samples with the same matrix element generator POWHEG-BOX and different MC showering models HERWIG ++ and PYTHIA 6, respectively. The parton shower uncertainty is compared with the nominal sample is shown in Fig. 5.20 for simulated $t\bar{t}$ events in each control and signal regions.

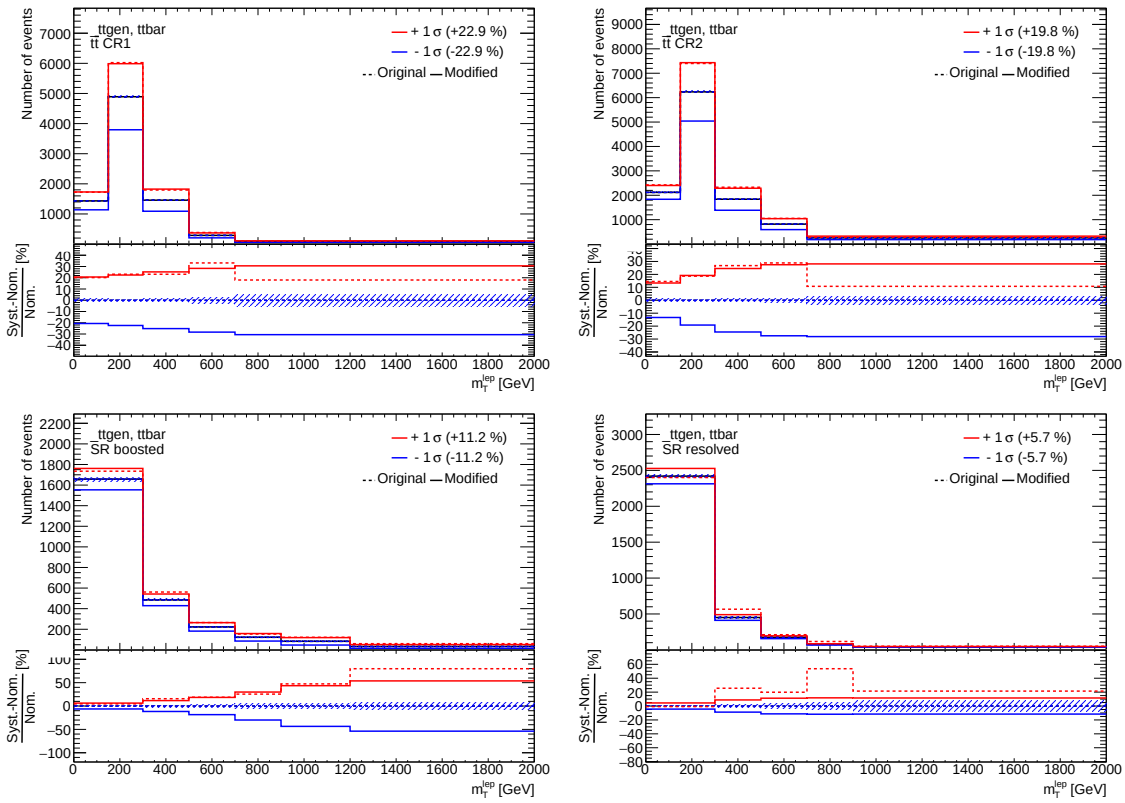

Figure 5.18: Matrix element generator uncertainty ('_ttgen') in $t\bar{t}$ events for the regions CR1 (top left), CR2 (top right), boosted signal region (bottom left), and resolved signal region (bottom right). 1σ up (red) and down (blue) variations of the respective systematic uncertainties with respect to the nominal sample (black). The original (dashed) and modified (solid) distributions of the respective systematic variation are shown. The hashed region corresponds to the statistical uncertainty per bin.

5.1.6 Statistical Analysis

A statistical analysis quantifies the compatibility between the measured data, expected SM background, and expected signal. Hypothesis testing is performed using a modified frequentist method as implemented in RooStats [106] and based on a profile likelihood that includes a treatment of the systematic uncertainties as nuisance parameters. A binned likelihood function $\mathcal{L}(\mu, \theta)$, is constructed as a product of Poisson probability terms over all

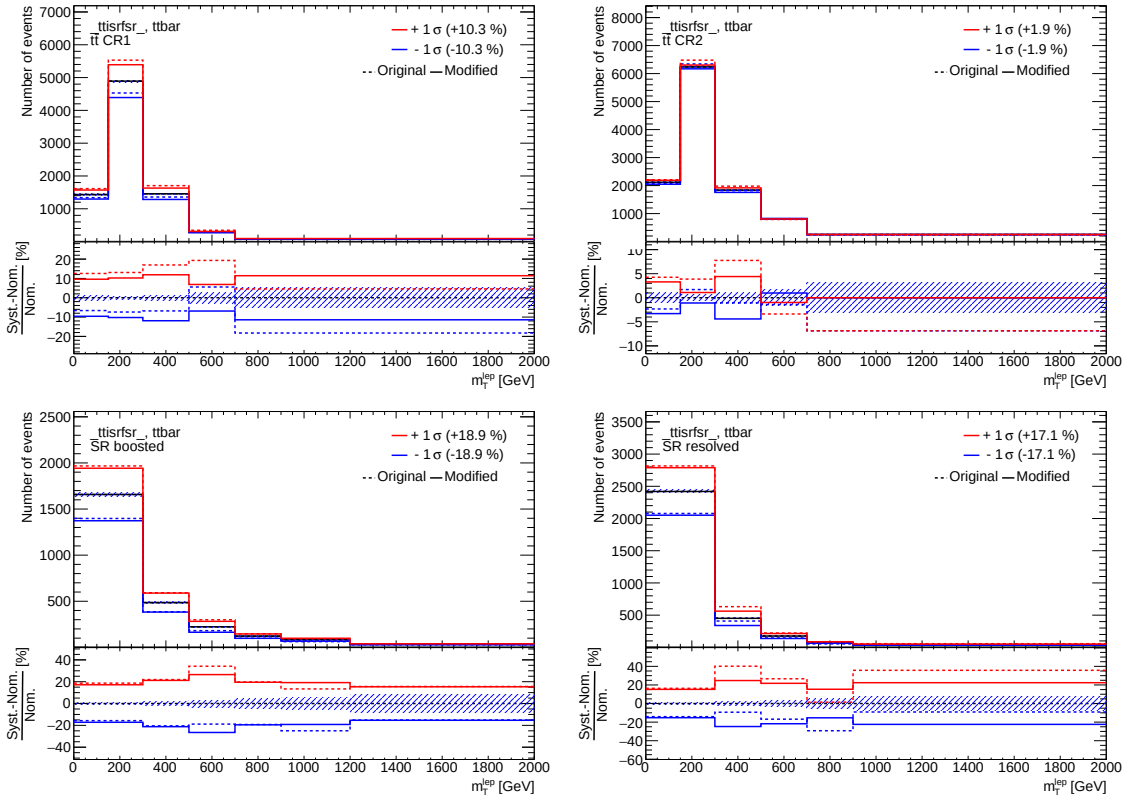

Figure 5.19: Initial and final state radiation uncertainties ('_ttisrfsr_') in $t\bar{t}$ events for the regions CR1 (top left), CR2 (top right), boosted signal region (bottom left), and resolved signal region (bottom right). 1σ up (red) and down (blue) variations of the respective systematic uncertainties with respect to the nominal sample (black). The original (dashed) and modified (solid) distributions of the respective systematic variation are shown. The hashed region corresponds to the statistical uncertainty per bin.

bins considered in the analysis for number of observed events N given the expected signal S and background events B (considering all backgrounds):

$$\mathcal{L}(\mu,\theta) = \prod_{i}^{n_{\text{bins}}} \left[\text{Pois}\left(N_i \middle| \mu S_i(\theta) + \sum_{b}^{n_{\text{bckgs}}} B_{i,b}(\theta) \right) \right] \times \prod_{i}^{n_\theta} \mathcal{M}(\theta_i, \bar{\theta}_i) \tag{5.11}$$

where μ is the multiplicative factor to the theoretical signal production cross section and θ represents the set of nuisance parameters. The number of expected events in a given bin

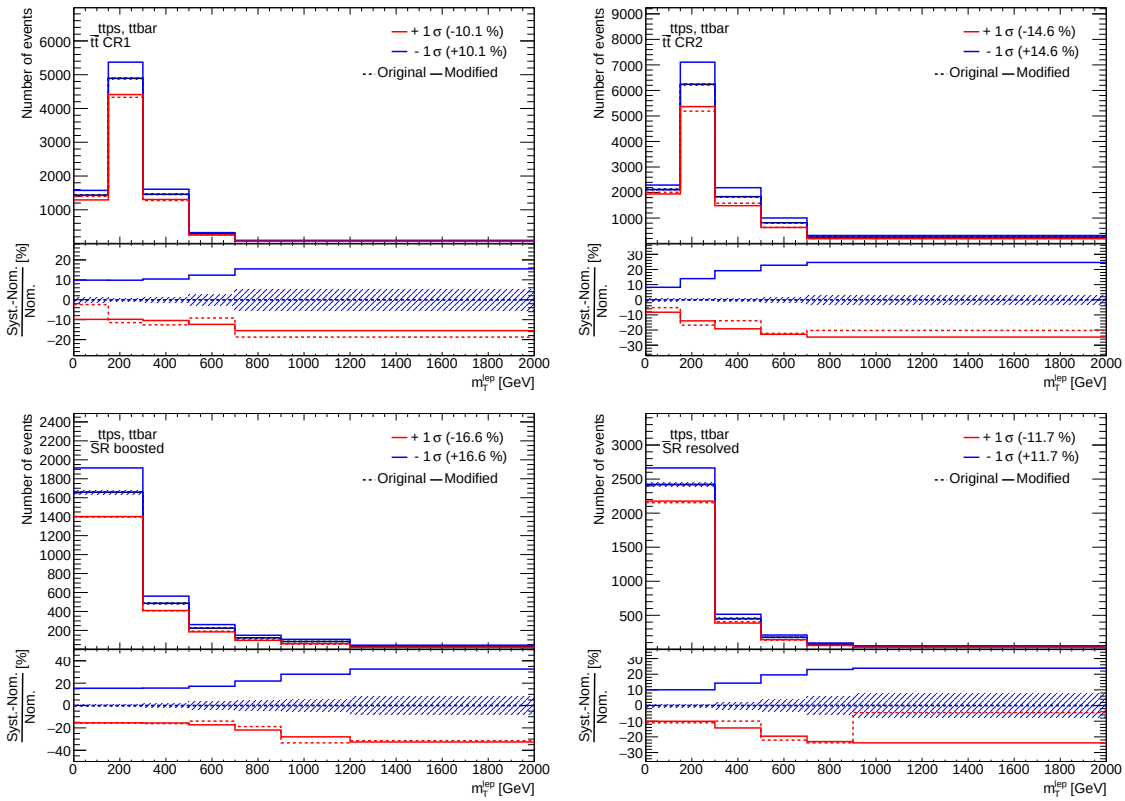

Figure 5.20: Parton shower uncertainty ('_ttps') in $t\bar{t}$ events for the regions CR1 (top left), CR2 (top right), boosted signal region (bottom left), and resolved signal region (bottom right). 1σ up (red) and down (blue) variations of the respective systematic uncertainties with respect to the nominal sample (black). The original (dashed) and modified (solid) distributions of the respective systematic variation are shown. The hashed region corresponds to the statistical uncertainty per bin.

depends on μ and θ. The nuisance parameters are given Gaussian ($\mathcal{M}_G$) or log-normal ($\mathcal{M}_{LN}$) priors in the likelihood fit:

$$\begin{aligned} \mathcal{M}_G(\theta, \bar{\theta}) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\theta - \bar{\theta})^2}{2\sigma^2} \right) \\ \mathcal{M}_{LN}(\theta, \bar{\theta}) &= \frac{1}{\sqrt{2\pi}\theta \ln(\kappa)} \exp\left(-\frac{\ln^2(\theta/\bar{\theta})}{2\ln^2 \kappa} \right) \end{aligned} \tag{5.12}$$

where $\bar{\theta}$ is the Gaussian mean (typically 0), σ the Gaussian variance (typically 1), and κ describes the distribution width in $\mathcal{M}_{LN}$. The log-normal prior is only used for uncertainties that predict a negative yield. During the fit, the parameters θ adjust the expectations for signal and background and the fitted values correspond to the best global agreement between data and prediction. Using this procedure, the impact from various systematic uncertainties

can be reduced by constraining them across the different bins. The parameter of interest in this analysis is the ratio between the fitted number of events and predicted events for the signal.

The test statistic q_μ, implemented with RooFit [106, 107], is defined as the profile likelihood ratio:

$$q_\mu = -2\ln\left(\frac{\mathcal{L}(\mu, \hat{\hat{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})}\right) \tag{5.13}$$

where $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximize the likelihood function (with the constraint $0 \leq \hat{\mu} \leq \mu$), and $\hat{\hat{\theta}}_\mu$ are the values of the nuisance parameters that maximize the likelihood function for a given value of μ. By setting $\mu = 0$ in the profile likelihood ratio, the test statistic (q_0) measures the compatibility of the observed data in the background-only hypothesis.

The p-value that quantifies the compatibility of the data with the background-only hypothesis, p_0, is estimated by integrating the distribution of q_0 from background-only pseudo-experiments, approximated using the asymptotic formulae [108], above the observed value of q_0. The estimation of p_0 is performed for each signal scenario. In the absence of any significant excess above the expected background, upper limits are placed on the signal production cross section using the $\mathrm{CL_s}$ method [109, 110]. Signal production cross sections are excluded at $\geq$95% confidence level if $\mathrm{CL_s} < 0.05$, where $\mathrm{CL_s}$ is computed using the asymptotic approximation [108],

The final fit variable for the baseline analysis is the reconstructed leptonic T mass (m_T^{lep}), as defined in Sect. 5.1.2. The fit is performed simultaneously in the two signal regions and two $t\bar{t}$ control regions, previously defined in Sect. 5.1.3. The binning of m_T^{lep} in each of these regions was optimized separately to ensure sufficient MC statistics in each bin, particularly for the dominant $t\bar{t}$ background.

5.1.6.1 Treatment of Nuisance Parameters

To improve the performance of the likelihood fit, nuisance parameters can be removed, symmetrized, or smoothed before the fitting procedure.

The normalization or shape components will be removed, i.e., pruned, if the impact on the nominal distribution of m_T^{lep} is less than 1% because they are found to negligibly impact the analysis result and their removal improves computation time. It is possible that both or neither of the components are pruned. Statistical uncertainties in each bin

of the discriminant distributions are also considered by explicit parameters in the fit. The systematic uncertainties that are not pruned are described in Sect. 5.1.5.

All one-sided systematic uncertainties are subjected to smoothing and symmetrization. This is most relevant for the $t\bar{t}$-modeling uncertainties on the parton shower and the choice of matrix element generator where only one variation is compared with the nominal prediction.

In addition, all double-sided systematics are symmetrized without any smoothing applied. The actual symmetrization is performed per bin taking the full absolute difference between up and down variation and applying this as a symmetric up and down variation around the nominal prediction within the bin. This step ensures that systematic variations in the likelihood fit do not have bins where both components have either more or less events than the nominal result. The nuisance parameter smoothing is achieved through a combination of histogram re-binning (bin merging) and reducing statistical fluctuations between bins [111].

5.1.7 Results

Before unblinding the data in the signal regions, expected limits were derived using the total expected background as pseudo-data. Deriving expected limits on $T\bar{T}$ production is critical to understand the behavior and impact of nuisance parameters in the likelihood fit. In addition, the $t\bar{t}$ control regions are fit separately and together to test the nuisance parameters' performance and validate the background estimation. After unblinding the data in the signal region, a simultaneous fit to the signal and control regions is performed to derive limits on $T\bar{T}$ production.

5.1.7.1 Fit to Asimov Data

The likelihood fit is performed using the total expected background as a pseudo-dataset ('Asimov' data) under the background-only hypothesis in all signal and control regions following the prescription in Sect. 5.1.6. Fitting to Asimov data offers a test of the systematic uncertainties, pulls and constraints, and to understand the expected sensitivity to a signal hypothesis. Uncertainties below the pruning threshold are omitted from the fit, as described in Sect. 5.1.6.1. The summary of this pruning step is illustrated in Fig. 5.23. The Asimov results are shown in Figs. 5.21 and 5.22.

The list of fitted nuisance parameters to the Asimov dataset are shown in Fig. 5.24. The fitted nuisance parameters are shown with their pulls and constraints compared to the pre-fit value of 0 and standard deviation of 1. A fitted standard deviation of 1 indicates the data does not have enough statistical power to reduce the original uncertainty. A pulled

nuisance parameter intimates the prediction needs to be shifted to improve agreement with the measured data.

After the Asimov fit, the constrained nuisance parameters are mostly those of the $t\bar{t}$-modeling uncertainties. As expected for the Asimov dataset, where the predicted yield and pseudo-data match, no pulls from the nominal value are observed.

As a further cross check of the fit results, the correlations between the nuisance parameters in the Asimov fit are shown in Fig. 5.25. Only nuisance parameters with a correlation coefficient of at least 10% with any other parameter are displayed. No significant correlations of nuisance parameters in the fit are observed, except for some stronger correlations between the $t\bar{t}$-modeling parameters with values of up to 40% and correlations between some large-R jet uncertainties. Note, the correlations are measured by their impact on a product over all fitted bins.

A ranking of the nuisance parameters according to the absolute post-fit effect on the fitted signal strength $\hat{\mu}$ in the fit to Asimov data for the signal-plus-background hypothesis can be seen in Fig. 5.26, where the signal hypothesis corresponds to $T\bar{T}$ ($m_T = 800$ GeV; BR($T \rightarrow Wb$)). After the nominal fit, the nuisance parameter ranking is obtained by rerunning the fit twice for each nuisance parameter. Each nuisance parameter is fixed to $\hat{\theta} \pm \sigma_\theta$, where the $\hat{\theta}$ (σ_θ) is the post-fit value (uncertainty), and a new value of μ is obtained for each fit $\mu'_\pm$. The difference between the nominal μ and the new value $\mu'_\pm$ represents the impact of the nuisance parameter on the signal strength. The uncertainty with the largest effect on μ are $t\bar{t}$ modeling uncertainties, which are also significantly constrained. Uncertainties on JES for large-R jets and b-tagging uncertainties are also significant.

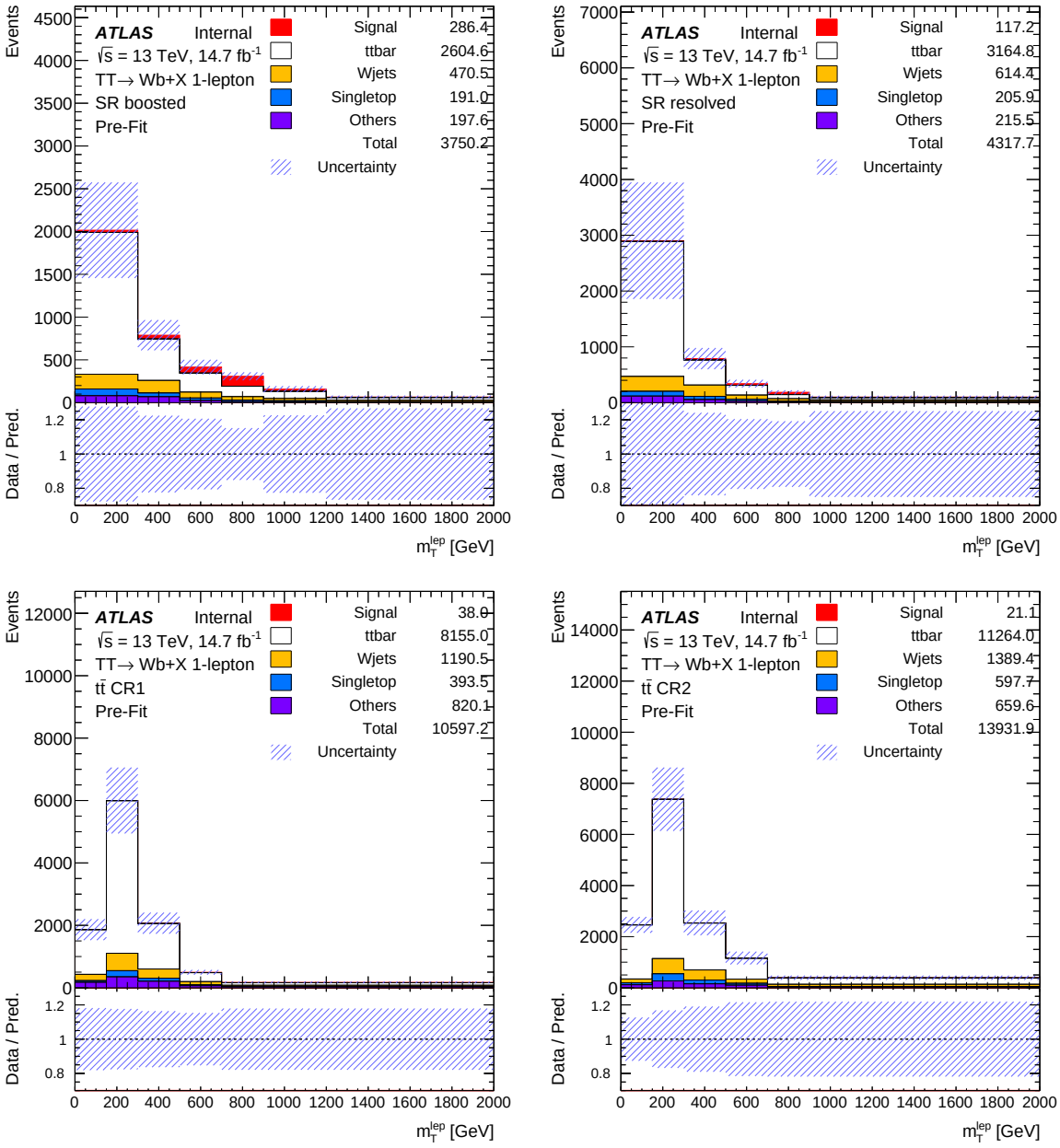

Figure 5.21: Comparison between Asimov data and prediction for the m_T^{lep} distribution before the fit to the Asimov data under the background-only hypothesis. The figures represent the boosted signal region (top left), the resolved signal region (top right), CR1 (bottom left), and CR2 (bottom right). The expected $T\bar{T}$ signal ($m_T = 800$ GeV; BR($T \to Wb$) = 1) is also shown. The hashed area represents the total uncertainty on the background.

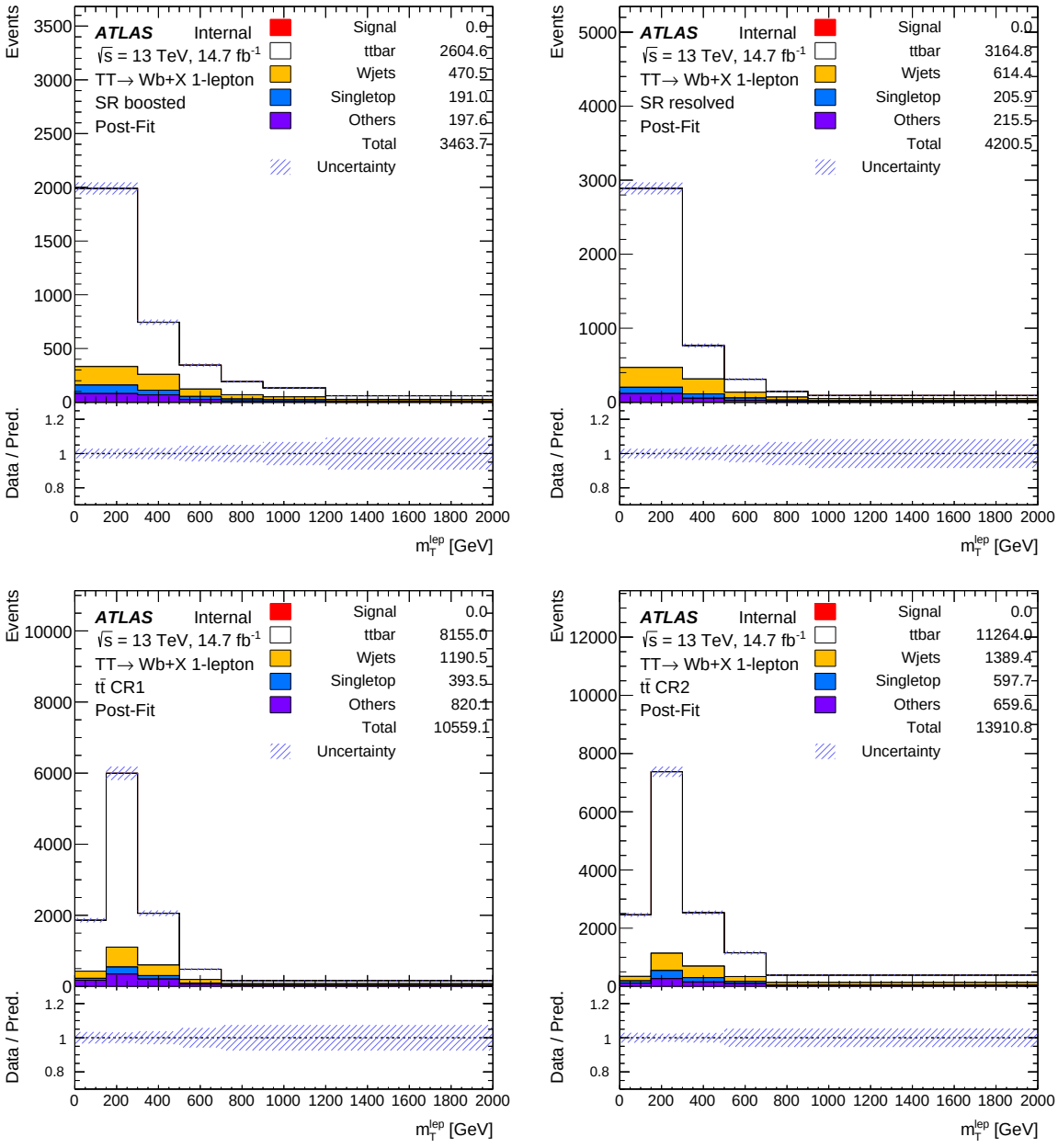

Figure 5.22: Comparison between Asimov data and prediction for the m_T^{lep} distribution after the fit to the Asimov data under the background-only hypothesis. The figures represent the boosted signal region (top left), the resolved signal region (top right), CR1 (bottom left), and CR2 (bottom right). The hashed area represents the total uncertainty on the background.

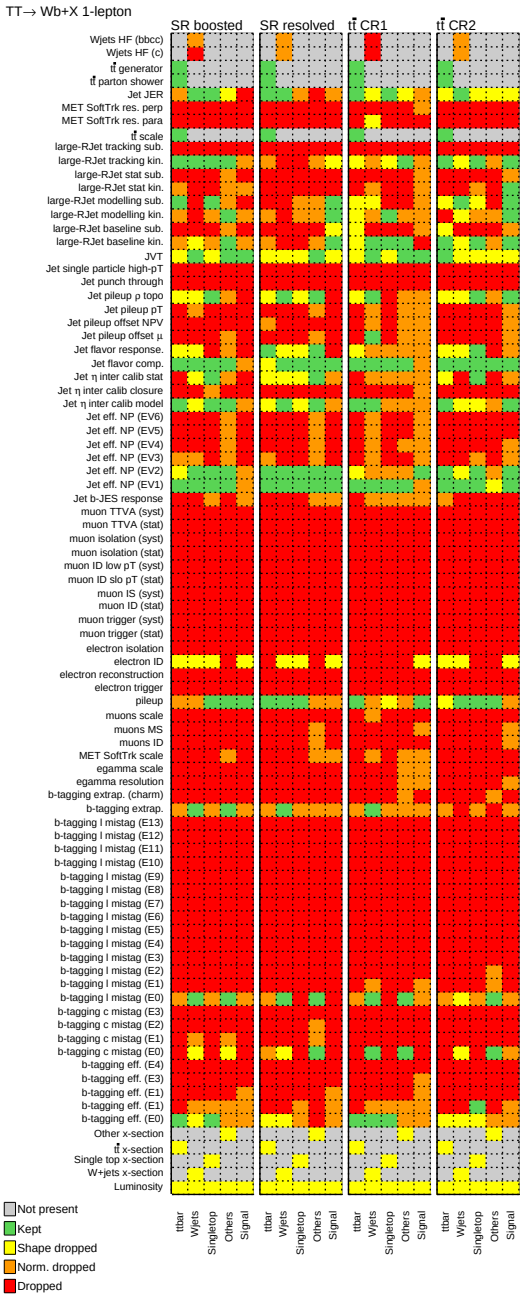

Figure 5.23: Summary of the pruning of the systematic uncertainties. The uncertainties that have the shape (normalization) component removed are yellow (orange) and those that are red were completely dropped from the fit. Uncertainties are green if both shape and normalization remain in the fit.

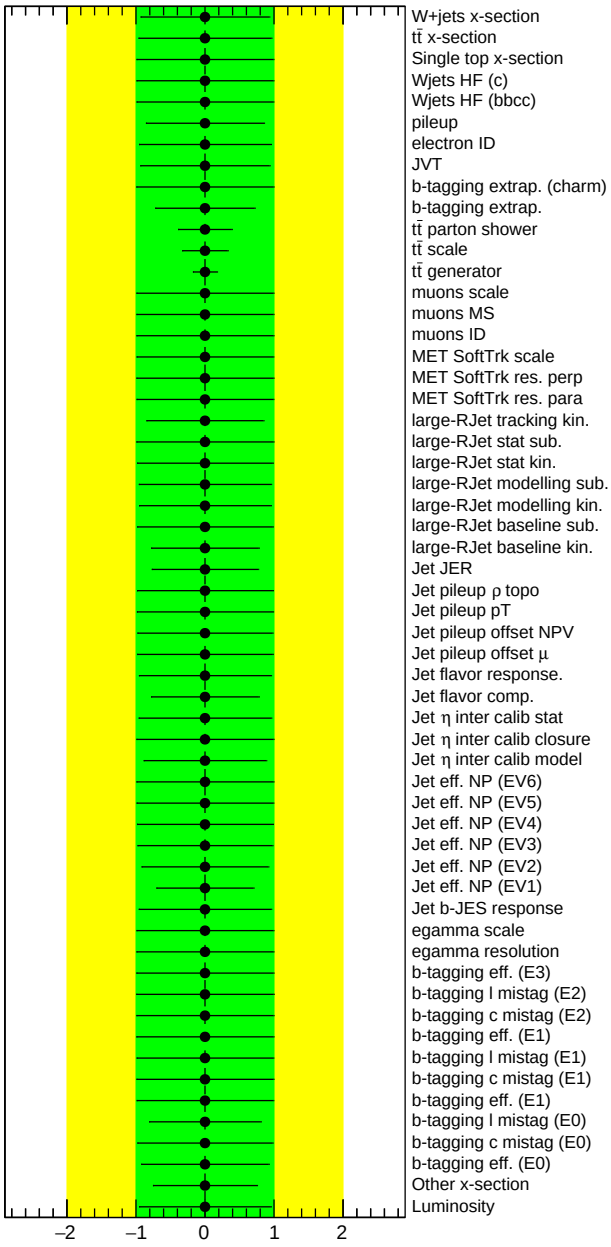

Figure 5.24: Fitted nuisance parameters in the background-only hypothesis to the Asimov dataset. Nuisance parameters are pulled if the fit deviates from 0 and constrained if the error is less than 1.

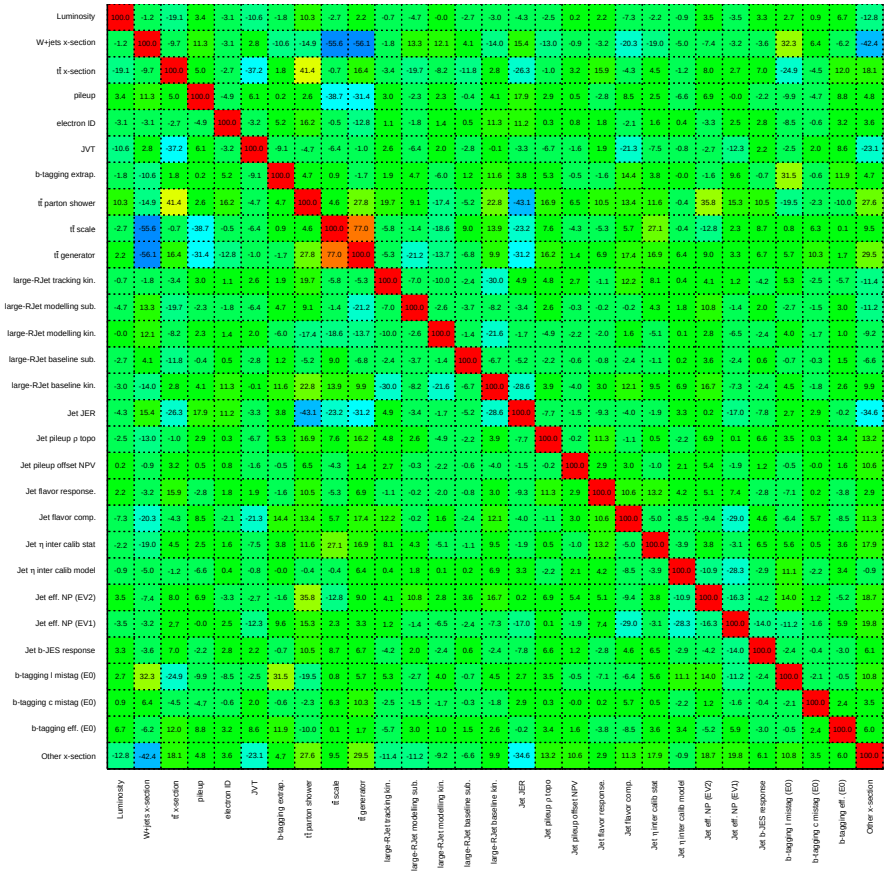

Figure 5.25: Correlation matrix showing the correlations of the nuisance parameters for the fit to the Asimov dataset in the signal+background hypothesis. The correlation scale legend corresponds to 100% (red), 0% (green) and -100% (violet). Only nuisance parameters with a correlation coefficient of at least 10% with any other parameter are displayed.

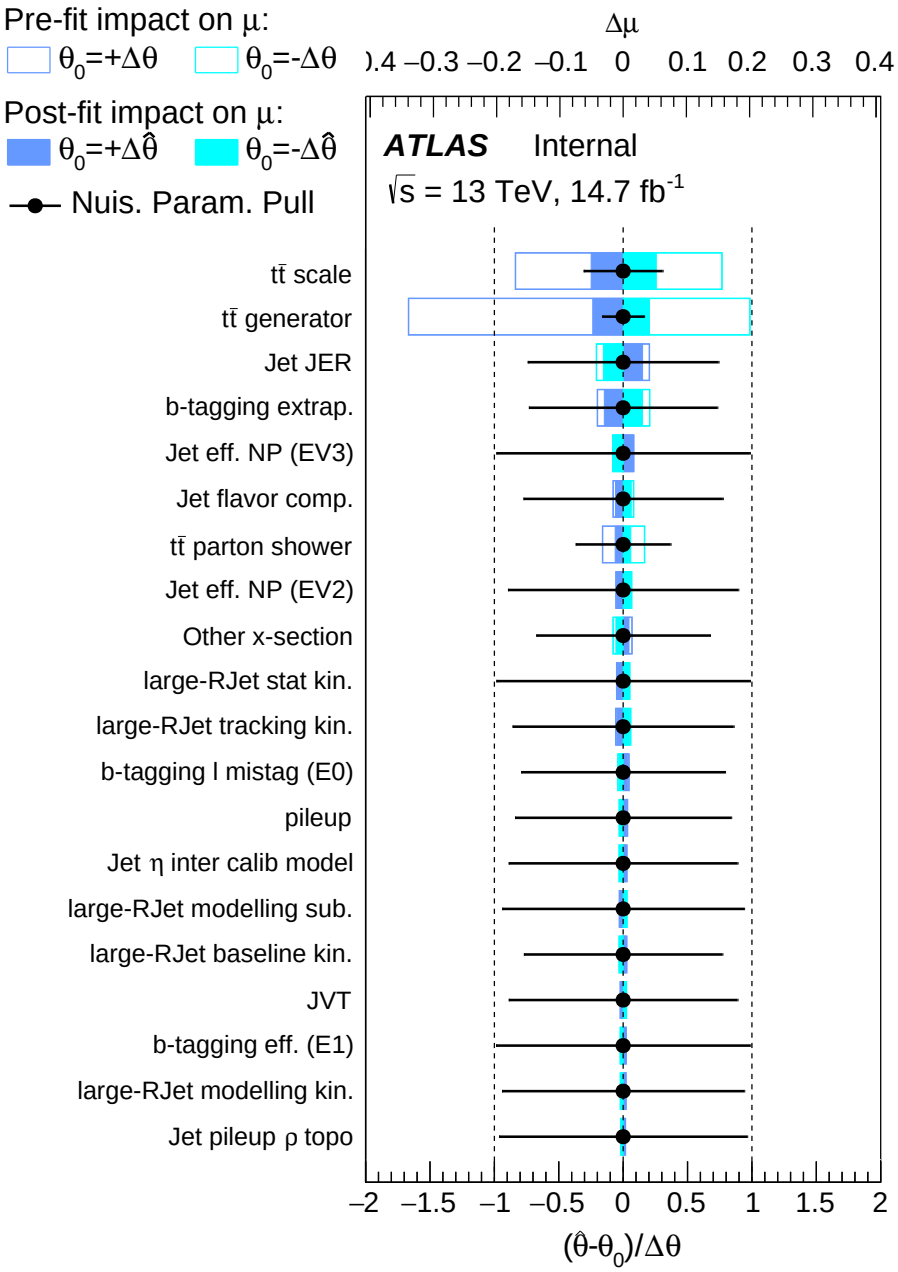

Figure 5.26: Ranking of nuisance parameters based on the fit to Asimov data in the signal-plus-background hypothesis according to their effect on the uncertainty on μ ($\Delta\mu$). The signal hypothesis corresponds to $T\bar{T}$ (m_T = 800 GeV; BR($T \to Wb$) = 1). The open boxes show the initial impact of each uncertainty on the precision of μ. The filled dark blue (light blue) area shows the impact on the measurement for each uncertainty after the profile likelihood fit at the +1 (−1) σ level (top axis). The black points and associated error bars show the fitted value of the nuisance parameters and their errors (bottom axis). Only the 20 highest ranked uncertainties on μ, ordered by impact from largest (top) to smallest (bottom), are shown.

Upper limits at the 95% confidence level (CL) on the $T\bar{T}$ production cross section are set for two benchmark scenarios as a function of m_T and compared to the theoretical prediction from `top++`, see Fig. 5.37. The resulting lower limit on m_T corresponds to the central value of the theoretical cross-section prediction. The scenarios considered involve different assumptions on the decay branching ratios. Under the assumption of $\mathrm{BR}(T \to Wb) = 1$, an expected 95% CL limit of $m_T = 980$ GeV is obtained. The corresponding expected 95% CL limit is $m_T = 872$ GeV for a weak-isospin singlet. This represents a significant improvement compared to run 1 searches [97,112], for which the observed 95% CL limits for the assumption of $\mathrm{BR}(T \to Wb) = 1$ were below 850 GeV. This result is competitive with a recent analysis performed by the CMS collaboration using data collected at $\sqrt{s} = 13$ TeV in 2015 which excluded masses below 876 GeV [15].

The same analysis is used to derive exclusion limits on vector-like T quark production for different values of m_T and as a function of $\mathrm{BR}(T \to Wb)$ and $\mathrm{BR}(T \to Ht)$. To probe the complete branching ratio plane spanned by both processes, the signal samples are re-weighted to the desired branching ratios and the complete analysis chain is repeated. Fig. 5.38 shows the expected T quark mass limits, alongside the observed limits, in the plane $\mathrm{BR}(T \to Ht)$ versus $\mathrm{BR}(T \to Wb)$.

5.1.7.2 Fit to Data in Control Regions

A fit on the $t\bar{t}$ control regions, individually and together, tests the constraints on nuisance parameters and the behavior of the likelihood fit using real data in a signal-depleted region. Fig. 5.27 shows the pre- and post-fit plots for fitting both $t\bar{t}$ control regions (CR1 and CR2) together and the fits to only CR1 and only CR2 are shown in Figs. 5.28 and 5.29, respectively.

The constraints, uncertainty less than 1, and pulls, mean different from 0, of the nuisance parameters (NPs) reveal structure in the uncertainties that is used to improve agreement between prediction and data. These are shown in the left plot in Fig. 5.30 with the fit to the combined two $t\bar{t}$ control regions compared to the fits obtained from each region individually. Few NPs exhibit pulls in opposite directions across the different fit configurations and all NPs are pulled less than 1σ, indicating the background estimation is in agreement with data prior to the profile likelihood fit. The most strongly pulled NPs are the $t\bar{t}$ modeling uncertainties.

5.1.7.3 Fit to Unblinded Data

After unblinding, the compatibility of the data with the background-only prediction is estimated using the same procedure as the fits to Asimov data described in Sect. 5.1.7.1.

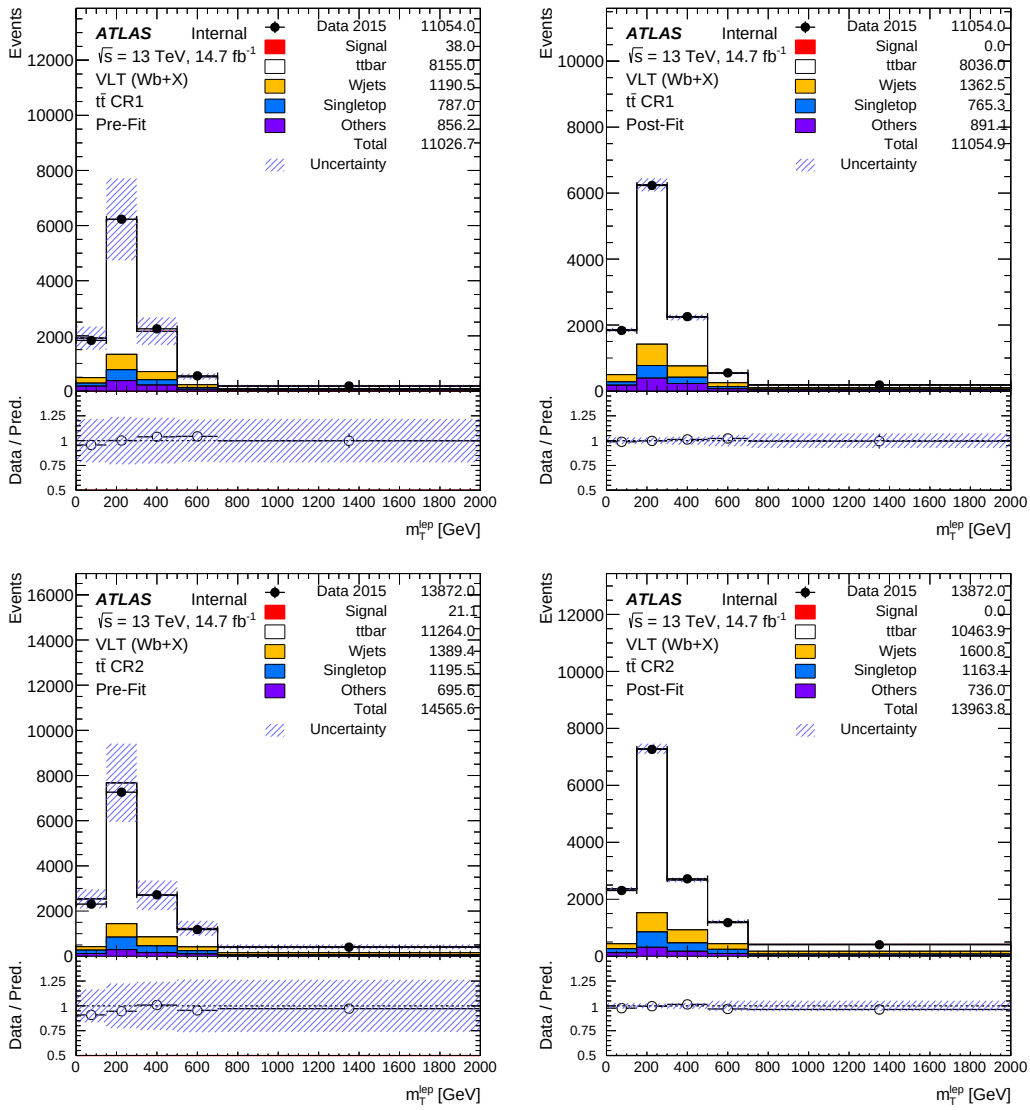

Figure 5.27: Comparison between data and prediction for the m_T^{lep} distribution before (left) and after (right) the simultaneous fit to both $t\bar{t}$ control regions under the background-only hypothesis. The top row shows CR1 and the bottom row shows CR2. The expected $T\bar{T}$ signal ($m_T = 800$ GeV; BR($T \to Wb$) = 1) is also shown. The hashed area represents the total uncertainty on the background.

Comparisons of the m_T^{lep} distributions in each region for both pre- and post-fit are shown Figs. 5.31 and 5.32 and summarized in Table 5.7. For illustration, the $m_T = 800$ GeV BR($T \to Wb$) = 1 signal model is shown overlaid on the total background prediction.

Fig. 5.33 displays the fitted nuisance parameters, which show similar constraints as expected from the fits to Asimov data and the data in the $t\bar{t}$ control regions. The corresponding correlation matrix for the fitted nuisance parameters are shown in Fig. 5.34. From Figs. 5.31 and 5.32, the agreement between the total prediction and data in the signal regions is not

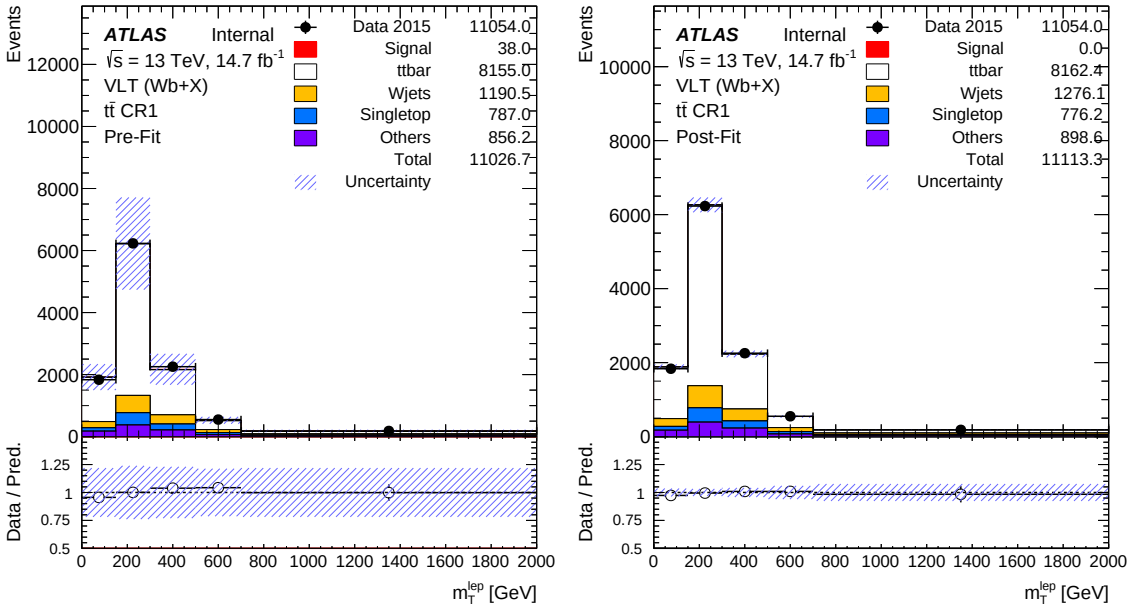

Figure 5.28: Comparison between data and prediction for the m_T^{lep} distribution before (left) and after (right) the fit to CR1 data under the background-only hypothesis. The expected $T\bar{T}$ signal ($m_T = 800$ GeV; BR($T \to Wb$) = 1) is also shown. The hashed area represents the total uncertainty on the background.

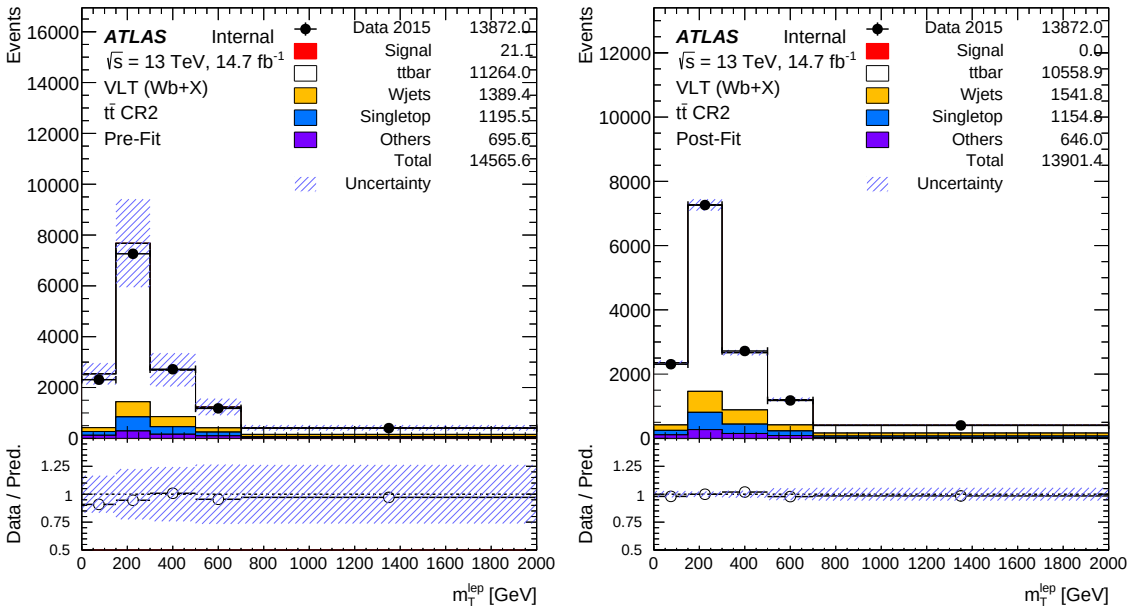

Figure 5.29: Comparison between data and prediction for the m_T^{lep} distribution before (left) and after (right) the fit to CR2 data under the background-only hypothesis. The expected $T\bar{T}$ signal ($m_T = 800$ GeV; BR($T \to Wb$) = 1) is also shown. The hashed area represents the total uncertainty on the background.

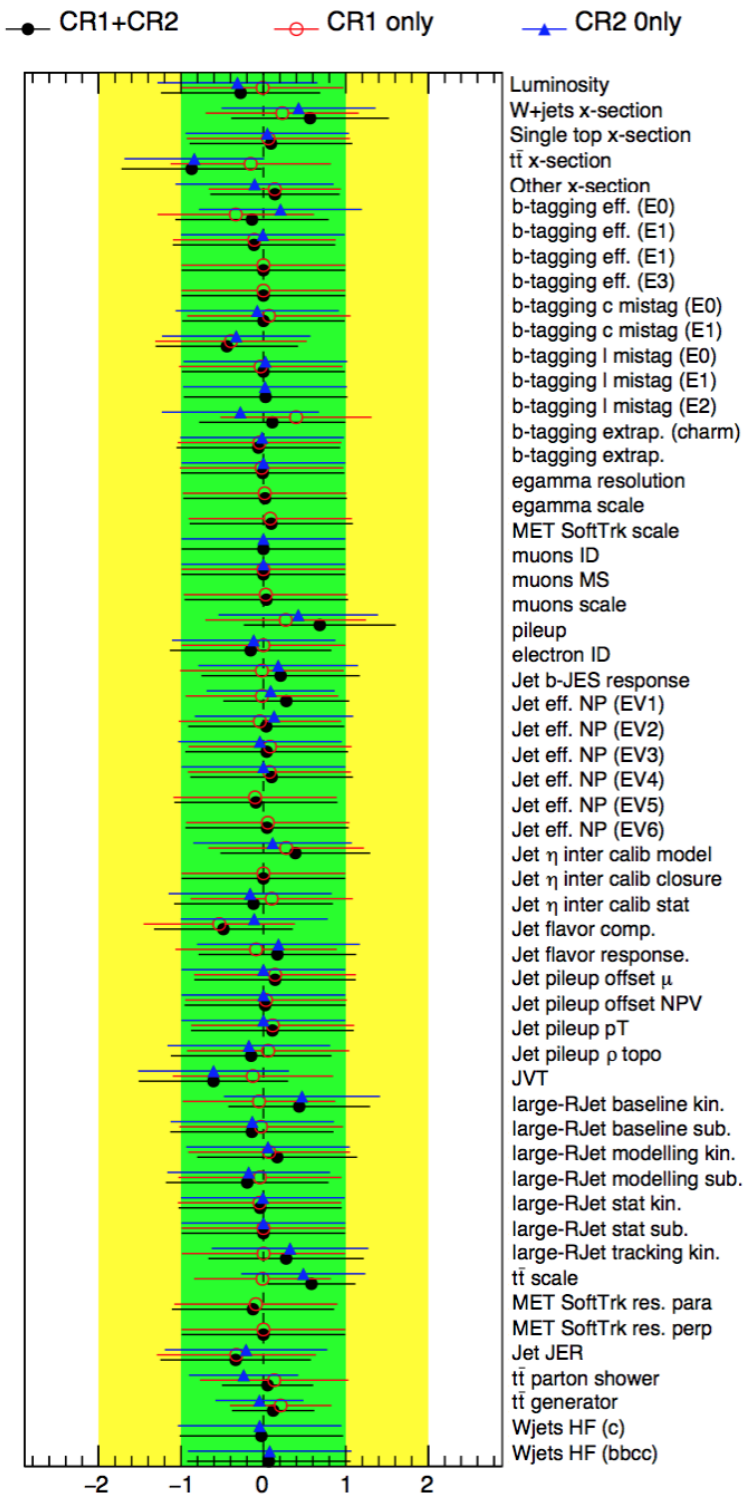

Figure 5.30: Fitted nuisance parameters for the background-only hypothesis in the fit to (CR1+CR2), CR1 only, and CR2 only as indicated in the legend.

Table 5.7: Event yields in all four signal and control regions after the fit to the background-only hypothesis. The uncertainties include statistical and systematic uncertainties. The uncertainties on the individual background components can be larger than the uncertainty on the sum of the backgrounds due to correlations.

	SR boosted	SR resolved	$t\bar{t}$ CR1	$t\bar{t}$ CR2
$t\bar{t}$	2130 ± 130	2370 ± 120	7710 ± 370	$10\,370 \pm 350$
W+jets	690 ± 110	880 ± 130	1860 ± 320	2160 ± 340
Single top	191 ± 17	202 ± 18	415 ± 35	628 ± 55
Others	232 ± 90	244 ± 98	90 ± 390	740 ± 300
Total Background	3246 ± 81	3690 ± 100	$10\,980 \pm 300$	$13\,900 \pm 290$
Data	3160	3777	11 054	13 872

as good as for both control regions. As a result of this disagreement, several nuisance parameters are pulled more significantly. The $t\bar{t}$ modeling uncertainties for the parton shower and the scale-variation are pulled both to lower the $t\bar{t}$ prediction within the first few bins, see Figs. 5.20 and 5.19.

In this context, the $t\bar{t}$ normalization is also pulled low, whereas the W+jets normalization is pulled high trying to accommodate the excess of data at high values of m_T^{lep}. This results in a slight excess in both signal regions between 500–900 GeV and a slight deficit above 900 GeV in m_T^{lep}.

A ranking of the nuisance parameters according to the absolute post-fit effect on the fitted signal strength $\hat{\mu}$ in the fit to data for the signal-plus-background hypothesis can be seen in Fig. 5.35, where the signal hypothesis corresponds to $T\bar{T}$ ($m_T = 800$ GeV; $\text{BR}(T \to Wb) = 1$). This procedure follows that used for the Asimov fit, see Sect. 5.1.7.1. The $t\bar{t}$ modeling uncertainties have the largest impact on μ and are also significantly constrained. Uncertainties on large-R jet JES and b-tagging uncertainties are also significant. The observed ranking is qualitatively very similar to that obtained from the Asimov dataset as shown in Fig. 5.26.

The p_0-value is computed for each signal scenario considered. No significant excess above the background expectation is found. A comparison of the post-fit agreement between data and prediction within both SRs in Figs. 5.31 and 5.32 show a slight excess of data in the range between 500–900 GeV of m_T^{lep} and a small deficit in data above 900 GeV. In this context, the observed upper limits on the $T\bar{T}$ production cross section are slightly weaker with respect to the expected sensitivity for low m_T, and conversely stronger for high m_T. The signal model corresponding to $m_T = 700$ GeV under the assumption of $\text{BR}(T \to Wb) = 1$ has the highest

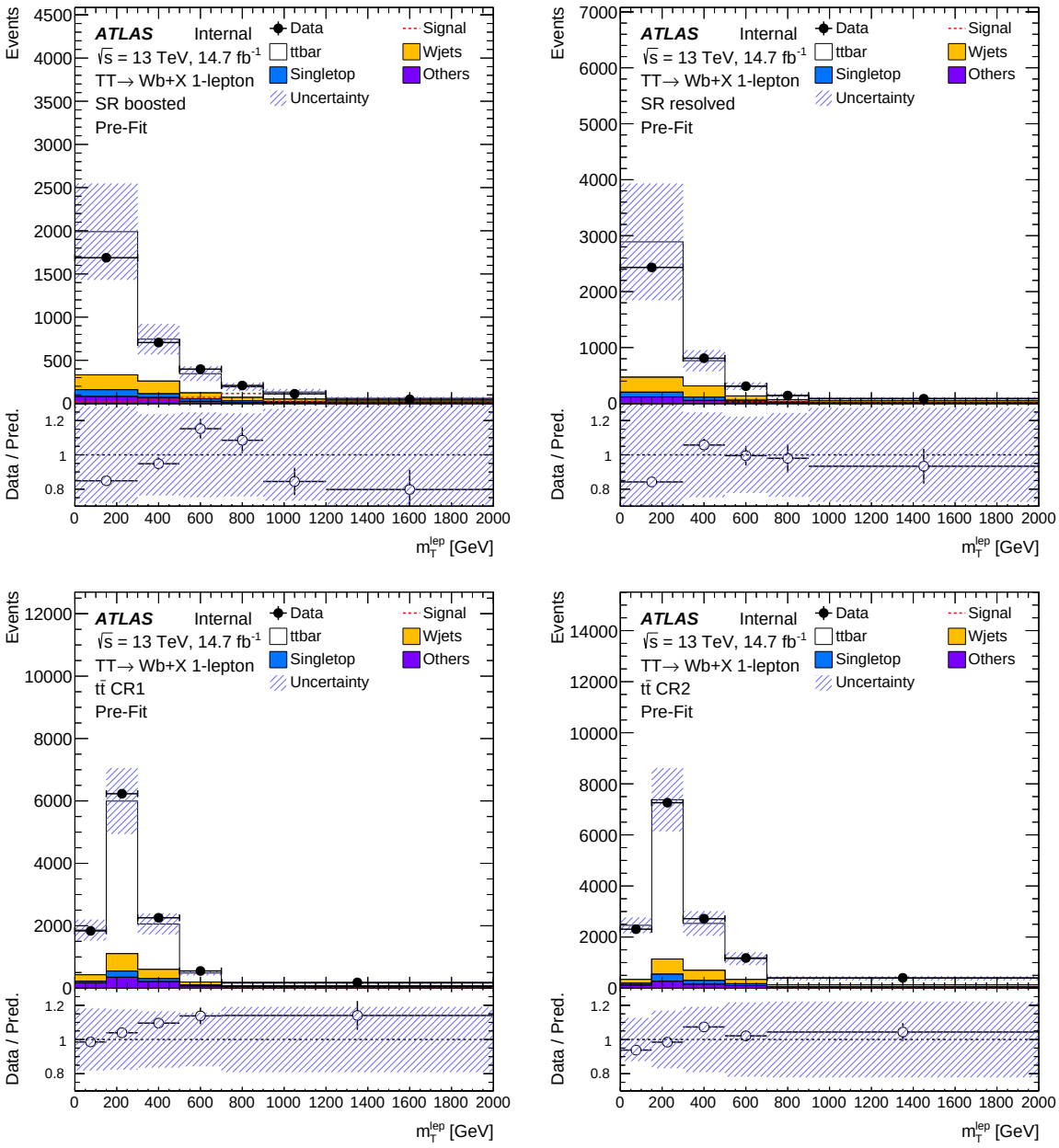

Figure 5.31: Comparison between data and prediction for the m_T^{lep} distribution before the fit to the data under the background-only hypothesis. The boosted signal region (top-left), resolved signal region (top-right), CR1 (bottom-left), and CR2 (bottom-right). The expected $T\bar{T}$ signal ($m_T = 800$ GeV; BR($T \to Wb$) = 1) is also shown. The hashed area represents the total uncertainty on the background.

local significance with 1.9σ standard deviations. A summary of the expected and observed significances for each signal scenario considered, BR($T \to Wb$) = 1 and SU(2) singlet, are shown in Fig. 5.36.

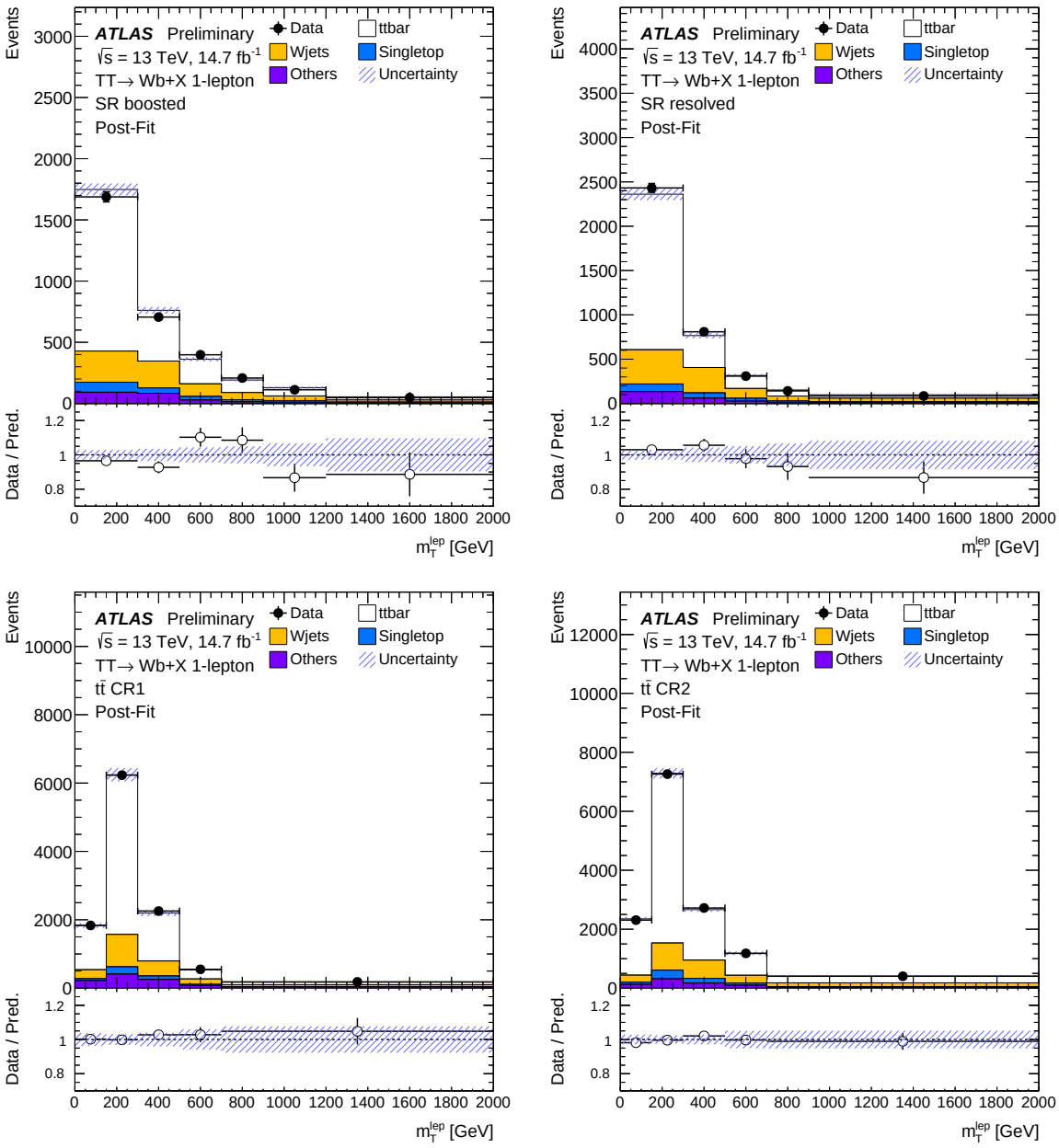

Figure 5.32: Comparison between data and prediction for the m_T^{lep} distribution after the fit to the data under the background-only hypothesis. The boosted signal region (top-left), resolved signal region (top-right), CR1 (bottom-left), and CR2 (bottom-right). The hashed area represents the total uncertainty on the background.

Upper limits at the 95% confidence level (CL) on the $T\bar{T}$ production cross section are set for two benchmark scenarios as a function of m_T and compared to the theoretical prediction from `top++` in Fig. 5.37. The resulting lower limit on m_T corresponds to the central value of the theoretical cross-section prediction. The scenarios considered involve different

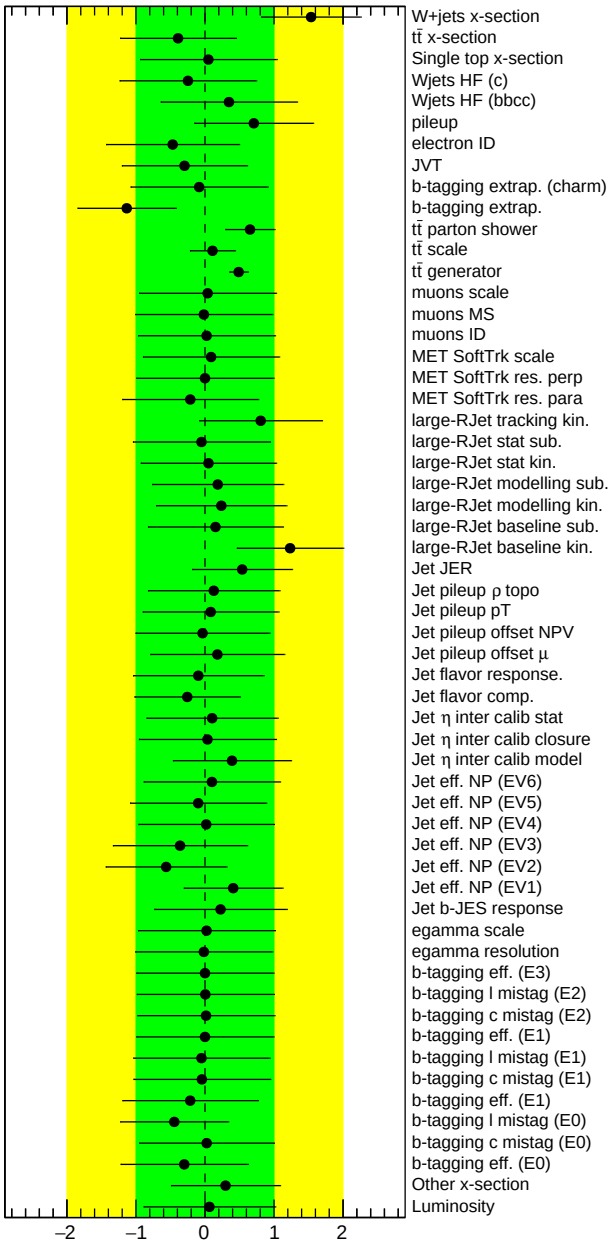

Figure 5.33: Fitted nuisance parameters in the background-only hypothesis to data.

assumptions on the decay branching ratios. Under the assumption of BR($T \to Wb$) = 1, an observed (expected) 95% CL limit of $m_T = 1090(980)$ GeV is obtained. The corresponding observed (expected) 95% CL limit is $m_T = 810(870)$ GeV for a weak-isospin singlet.

The same analysis is used to derive exclusion limits on vector-like T quark production for different values of m_T and as a function of BR($T \to Wb$) and BR($T \to Ht$). Fig. 5.38

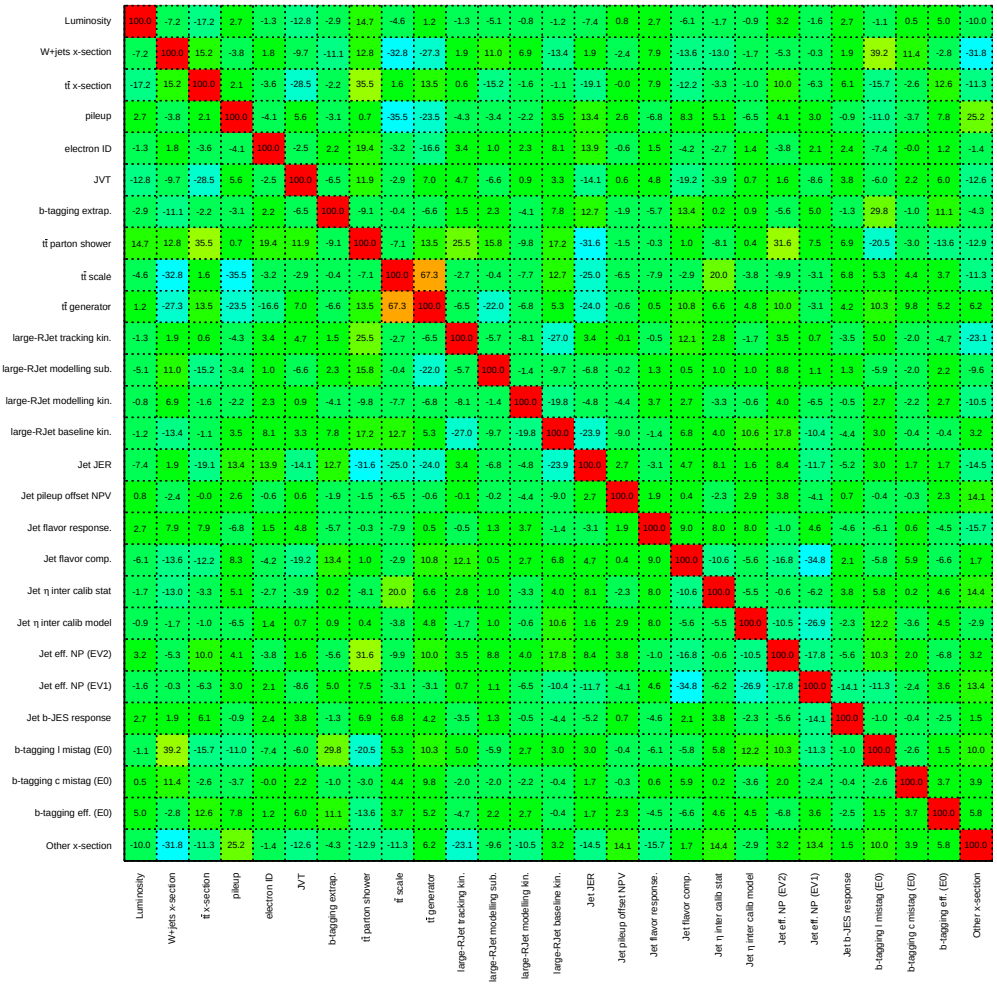

Figure 5.34: Correlation matrix showing the correlations of the nuisance parameters for the fit to data in the background-only hypothesis. The correlation scale legend corresponds to 100% (red), 0% (green) and -100% (violet). Only nuisance parameters with a correlation coefficient of at least 10% with any other parameter are displayed.

shows the expected and observed T quark mass limits in the plane BR($T \to Ht$) versus BR($T \to Wb$), obtained by linear interpolation of the calculated $\mathrm{CL_s}$ versus m_T.

5.1.8 Conclusion

A search for pair production of vector-like quarks in the lepton+jets final state has been presented using 14.7 fb^{-1} of integrated luminosity collected by the ATLAS detector in 2015 and 2016. The analysis targets final states in which the VLQ has significant decays to a high–p_{T} W boson and b-quark, selecting events with one charged lepton, high missing transverse momentum, at least two b-tagged jets, and either one large-radius or two small-radius jets

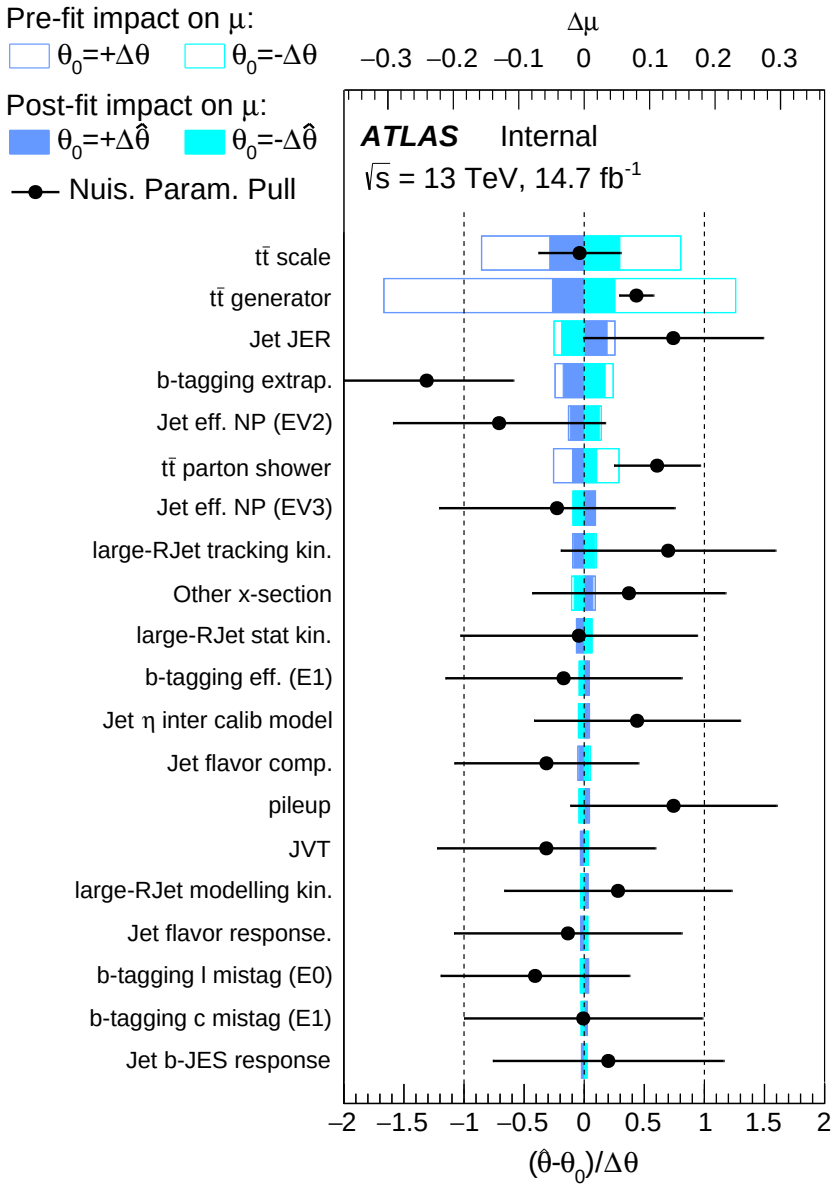

Figure 5.35: Ranking of nuisance parameters based on the fit to data in the signal-plus background hypothesis according to their effect on the uncertainty on μ ($\Delta\mu$). The signal hypothesis corresponds to $T\bar{T}$ ($m_T = 800$ GeV; BR($T \to Wb$) = 1). The open boxes show the initial impact of each uncertainty on the precision of μ. The filled dark blue (light blue) area shows the impact on the measurement for each uncertainty after the profile likelihood fit at the $+1$ (-1) σ level (top axis). The black points and associated error bars show the fitted value of the nuisance parameters and their errors (bottom axis). Only the 20 highest ranked uncertainties on μ, ordered by impact from largest (top) to smallest (bottom), are shown.

consistent with a boosted or resolved hadronically decaying W boson, respectively. The analysis also exploits angular correlations and the high scalar sum of transverse momenta

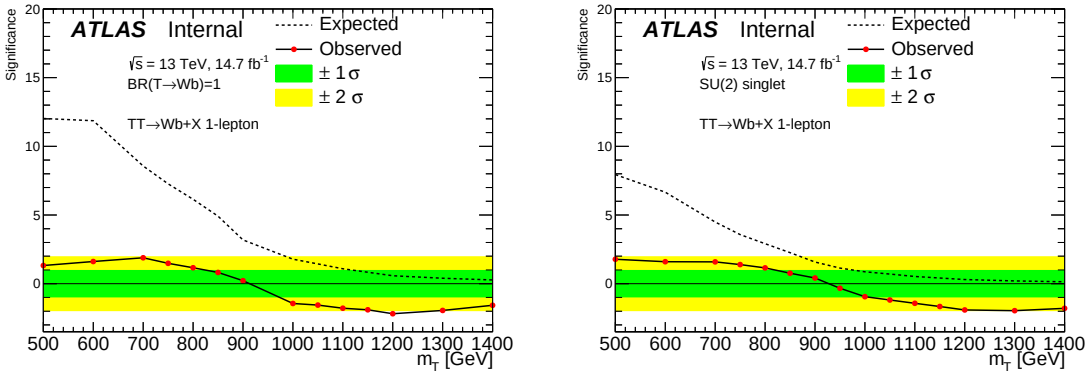

Figure 5.36: Expected (dashed black line) and observed (solid black line) significance as a function of T quark mass under the assumption of BR($T \to Wb$) = 1 (left) and SU(2) singlet (right). Yellow (green) bands for ± 1 (2) σ are also drawn.

expected for the VLQ signal hypothesis. The reconstructed mass of the leptonically decaying VLQ is used as the final discriminant.

The observed data does not have a significant excess compared to the background estimation. Lower limits on the mass of a VLQ for all possible combinations of branching ratios depict the sensitivity of the analysis to the $T \to Wb$ decay mode. The assumption that the W boson is produced in the VLQ decay and decays leptonically reduces sensitivity to the $T \to Ht/Zt$ decay modes, where the lepton is produced from a non-prompt W boson or Z boson. Due to the modest excess of events between $500 < m_T^{\text{lep}} < 900$ GeV, the analysis loses sensitivity to low mass VLQs for the $T \to Ht/Zt$ decay modes. Similarly, the deficit of events at large values of m_T^{lep} results in stronger limits for BR($T \to Wb$) = 1.

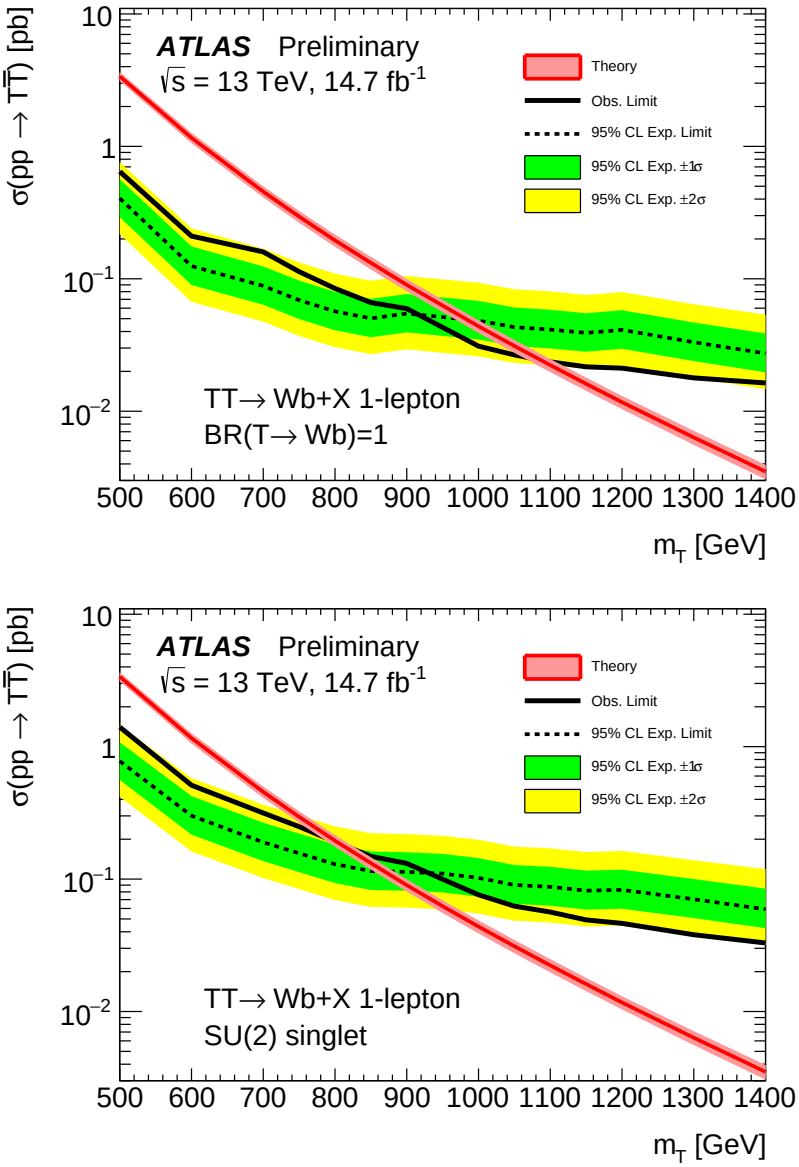

Figure 5.37: Expected (dashed black line) and observed (solid black line) upper limits at the 95% CL on the $T\bar{T}$ cross section as a function of m_T under the assumption of BR($T \to Wb$) = 1 (top) and SU(2) singlet (bottom). The surrounding green and yellow bands correspond to ± 1 and ± 2 standard deviations, respectively, around the expected limit. The thin red line and band show the theoretical prediction and its ± 1 standard deviation uncertainty.

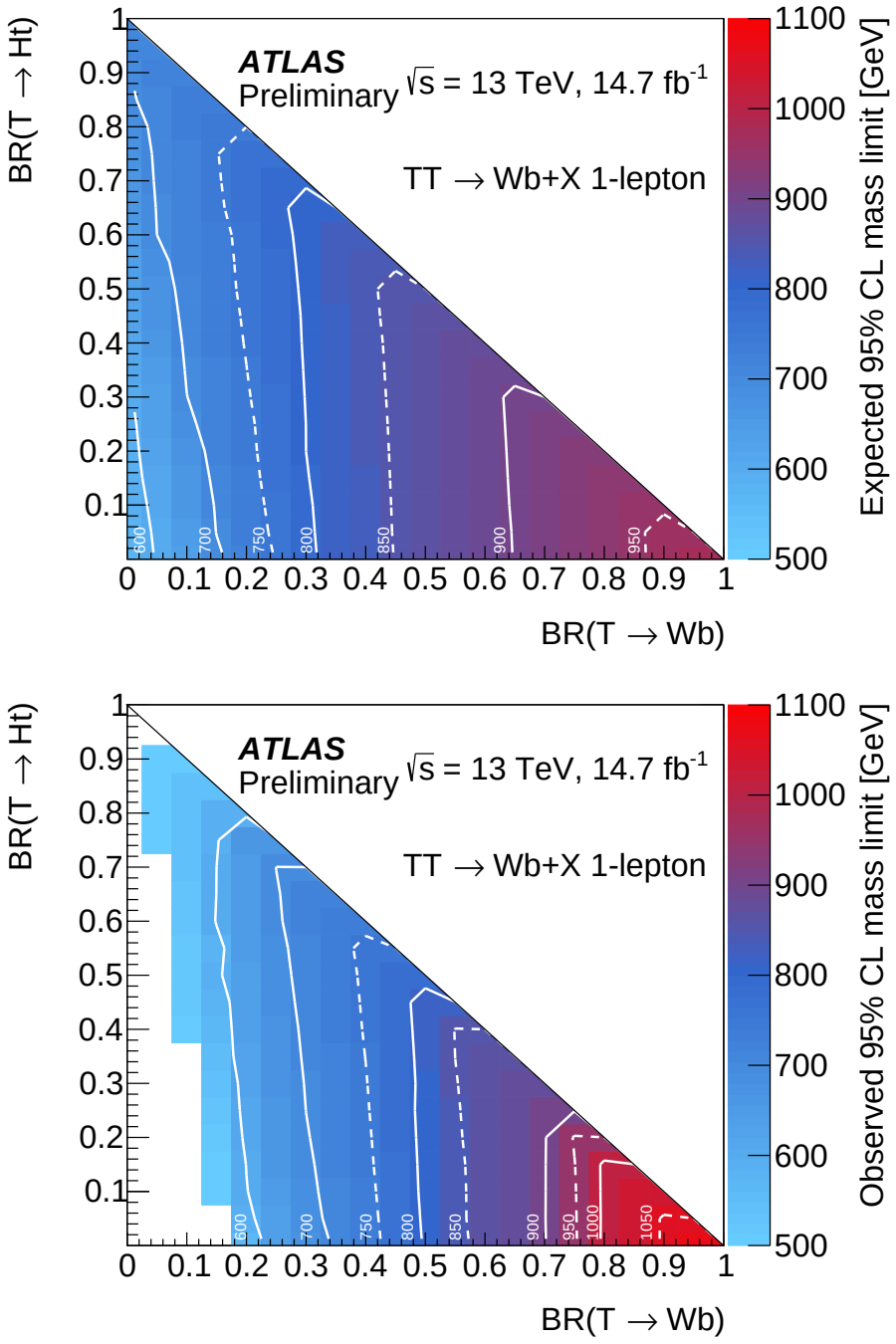

Figure 5.38: Expected (top) and observed (bottom) 95% CL upper limits on the mass of the T quark in the branching ratio plane of BR($T \to Wb$) versus BR($T \to Ht$). Contour lines show respective mass limits across the plane.

5.2 $T\bar{T}/B\bar{B} \rightarrow$ jets

The VLQ search previously described targets a dominant branching ratio $T \rightarrow Wb$ in the one lepton final state. While the lepton offers a 'clean' detector signature on which to trigger and easily reject QCD multi-jet events, the analysis is insensitive to the dominant branching ratio to all-hadronic final states. In the event of a discovery, it will be important to measure different branching ratios to identify what kind of VLQ has been observed. To complement the existing analyses, the all-hadronic final state is targeted in the following analysis. This final state has not been explored by the ATLAS or CMS collaborations for pair produced VLQs, providing a unique opportunity to search for VLQs in a new region of phase space.

The analysis strategy is optimized assuming the pair production of VLQs decaying to $T\bar{T} \rightarrow Wb/Zt/Ht$ or $B\bar{B} \rightarrow Wt/Zb/Hb$. The QCD multi-jet background, which dominates fully hadronic final states at hadron colliders, is reduced by requiring multiple high–p_{T} and b–tagged small-R jets. A variable–R re-clustering algorithm combines small-R jets to reconstruct decays of heavy objects which are then classified by a neural network as either an electroweak boson ($V \equiv W/Z$), Higgs boson (H), top quark, or 'light' jet. Events are categorized based on the number of V-tags, H-tags, top-tags, and b-tags. To be orthogonal with an existing high–$E_{\mathrm{T}}^{\mathrm{miss}}$ search [2], events are analyzed if $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV. Similar to the $T\bar{T} \rightarrow Wb + X$ analysis described previously in Sect. 5.1, a profile likelihood fit is performed to extract the excluded cross section as a function of m_{VLQ} and branching ratios.

5.2.1 Data and Simulated Samples

The search presented here uses the data collected by the ATLAS detector in 2015 and 2016, corresponding to an integrated luminosity of 36.1±1.2 fb^{-1}. Only collision data collected when all detector components are fully functional are considered for this analysis.

5.2.1.1 Signal production

The MC samples for $T\bar{T}$ production described in Sect. 5.1.1.1 are also used in this analysis. Additionally, the analysis uses simulated $B\bar{B}$ events generated using the same configuration as the $T\bar{T}$ samples. $B\bar{B}$ events are produced in the SU(2) singlet model, but are forced to decay with a branching ratio of 1/3 to each of the three modes (Wt, Zb, and Hb). To achieve a specific branching ratio, the same re-weighting procedure described for $T\bar{T}$ events is applicable to the $B\bar{B}$ events. The $B\bar{B}$ samples are produced at the same mass points and cross sections as the $T\bar{T}$ samples.

5.2.1.2 Background production

The dominant SM backgrounds considered by this analysis are multi-jet, $t\bar{t}$, and single top. The multi-jet background is estimated using a data-driven technique, see Sect. 5.2.4. However, simulated multi-jet events are used to study distributions when a data-driven estimate is not possible. The simulated multi-jet events are produced with PYTHIA 8.186 [79] using the A14 [80] set of tuned parameters for the underlying event and the leading-order NNPDF2.3 [84] PDFs. The renormalization and factorization scales are set to the average p_{T} of the two leading jets.

Simulated SM $t\bar{t}$ events with zero leptons are generated with the same setup and parameters as the one-lepton filter sample Sect. 5.1.1.2. Additionally, the $t\bar{t}$ sample with a one-lepton filter is also used in this analysis to estimate sensitivity to $t\bar{t}$ events with a real lepton that pass the event selection.

Finally, single top production, $Wt-$ and $t-$channels, is also generated by POWHEG-BOX +PYTHIA 6, using the `Perugia2012 tune` for the showering and the CT10 PDF set.

The $t\bar{t}+V$, $V = W/Z$, background is modeled using samples produced with MADGRAPH +PYTHIA 8 [90]. Similarly, the $t\bar{t}+H$ background is modeled using `aMC@NLO`+PYTHIA 8. The production of two τ leptons in association with $t\bar{t}$ events is modeled using MADGRAPH +PYTHIA 8. The $t\bar{t}+V$, $t\bar{t}+H$, and $t\bar{t}+\tau\tau$ contribute a very small portion of the background to this analysis and are collectively referred to as $t\bar{t}+X$.

5.2.2 Event Selection

Events are initially selected for the analysis if they pass an unprescaled jet-based trigger. The analysis requires special consideration of which triggers to use because the jet-based triggers are topology-dependent. Any triggers in the analysis must be accompanied by an event selection that ensures the analysis is not sensitive to any mis-modeling, i.e., the trigger efficiency must be $> 99\%$ in the analysis region. This ensures that no additional corrections and uncertainties due to data and MC mis-modeling need to be introduced in the analysis. To choose the appropriate triggers for the analysis, trigger efficiency is studied in simulated $T\bar{T}$ events. The most performant triggers are then studied in data and simulated multi-jet events to determine the optimal combination of triggers and offline selection that will be required.

The set of triggers tested for the analysis are based on H_{T} (`HT700` and `HT1000`) and re-clustered (RC) jet p_{T} (`j360_a10r` and `j400_a10r`). The H_{T} is calculated as the sum of the p_{T} of all small-R jets reconstructed in the last chain of the trigger. Small-R jets formed in the last chain of the trigger are re-clustered within a radius of 1.0 form the RC jets used

in the trigger. Triggers names that contain L1J100 (L1J75) require a jet identified with the level-1 trigger, reconstructed in 8×8 grid in the calorimeter, with a p_{T} above 100 (75) GeV.

The overall acceptance, the fraction of accepted out of total events, and full-efficiency acceptance, fraction of accepted events above 99% efficiency out of the total number events, are shown for $m_T = 1$ TeV events in Tables 5.8 and 5.9. The corresponding signal acceptances are shown in Figs. 5.39 and 5.40 as a function of the sum of all small-R jet p_{T} (H_{T}) and large-R jet p_{T}. For signal events, combining the RC jet trigger and the H_{T} trigger improves the acceptance for full-efficiency.

The comparison between data and multi-jet MC efficiencies are shown in 5.41. All simulated events are included in the efficiency calculation and data events are only considered if the event was collected during a period when the detector was fully operationally. The RC jet trigger does not improve the acceptance at full efficiency in data or multi-jet MC. For simplicity, the H_{T} trigger HT1000_L1J100 is the only trigger implemented in this analysis. Furthermore, neither large-R jets nor fixed–radius RC jets, which are needed for using the RC jet trigger, are used in this analysis. An offline cut of H_{T}> 1150 GeV for 2015 and 2016 events ensures full efficiency of the H_{T} triggers. The same offline H_{T} threshold is applied to 2015 and 2016, despite the lower thresholds available in the 2015 dataset, to maintain a similar kinematic phase space from both datasets.

Table 5.8: 2015 triggers tested in simulated $T\bar{T}$ ($m_T = 1$ TeV) events with the overall and full-efficiency acceptances relative to H_{T} and large-R jet p_{T}.

HLT Trigger	Acceptance (%) Overall	Full-Efficiency
H_{T}		
HT700_L1J75	99	96
j360_a10r_L1J100	97	75
HT700_L1J75 OR j360_a10r_L1J100	99	97
Leading large-R jet p_{T}		
HT700_L1J75	99	88
j360_a10r_L1J100	97	93
HT700_L1J75 OR j360_a10r_L1J100	99	96

After the trigger selection, data quality requirements are imposed on the events to ensure all detector components were operating correctly and that at least one primary vertex was reconstructed. Physics objects are reconstructed according to Sect. 4.2 and then undergo an overlap removal procedure to avoid double-counting energy in the detector. The working

Table 5.9: 2016 triggers tested in simulated $T\bar{T}$ ($m_T = 1$ TeV) events with the overall and full-efficiency acceptances relative to H_{T} and large-R jet p_{T}.

HLT Trigger	Acceptance (%) Overall	Full-Efficiency
H_{T}		
`HT1000_L1J100`	94	87
`j360_a10r_L1J100`	95	65
`HT1000_L1J100 OR j360_a10r_L1J100`	97	89
Leading large-R jet p_{T}		
`HT1000_L1J100`	94	45
`j400_a10r_L1J100`	95	88
`HT1000_L1J100 OR j400_a10r_L1J100`	97	88

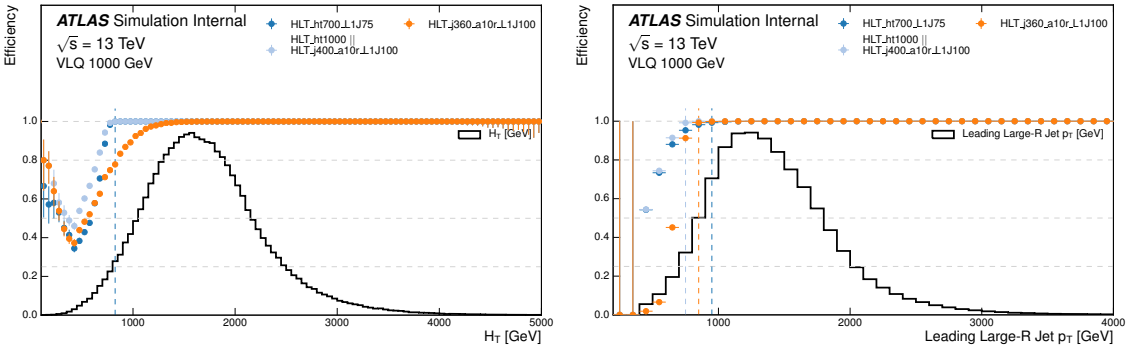

Figure 5.39: Trigger efficiencies as a function of H_{T} (left) and leading large-R jet p_{T} (right). The 2015 H_{T} trigger, large-R jet trigger, and a logical OR of both are drawn. The vertical dashed lines shown the value at which point the efficiency becomes >99%. The black histograms display the underlying distribution from the VLQ events.

points for the analysis physics objects are harmonized with other VLQ searches in the ATLAS collaboration using the full 2015 and 2016 dataset. The lepton working points were optimized in other VLQ analyses and incorporated here. Electrons are identified with the tight likelihood working point and the isolation $p_T^{\mathrm{varcone}}(R_{\mathrm{max}} = 0.2)$ (≡'`FixedCutTightTrackOnly` isolation') with $I = 0.06$. Muons are identified with the medium working point and the `FixedCutTightTrackOnly` isolation with $R_{\mathrm{max}} = 0.3$.

All small-R jets must have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. If the jet is centrally-located ($|\eta| < 2.4$) and has low–p_{T} ($p_{\mathrm{T}} < 60$ GeV), it must have a JVT output > 0.59 to be used in the analysis. A small-R jet is considered b-tagged if it satisfies the 70% working point of the `mv2c10` algorithm. The 70% working point is not harmonized with the other VLQ

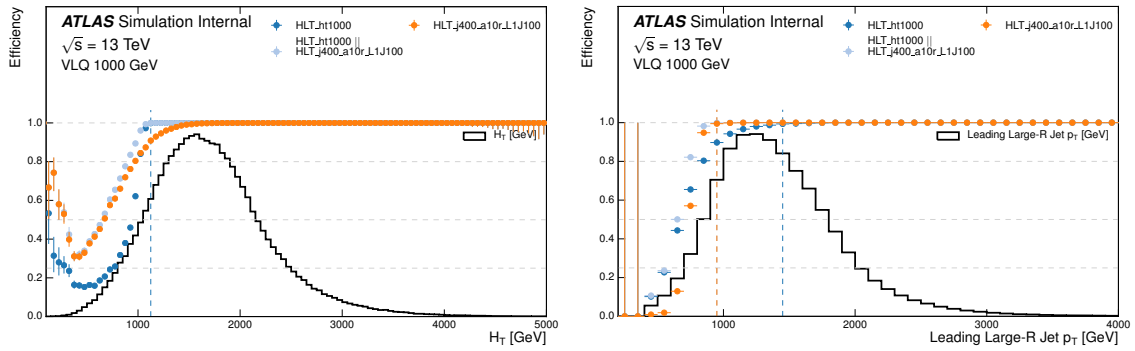

Figure 5.40: Trigger efficiencies as a function of H_{T} (left) and leading large-R jet p_{T} (right). The 2016 H_{T} trigger, large-R jet trigger, and a logical OR of both are drawn. The vertical dashed lines show the value at which point the efficiency becomes $> 99\%$. The black histograms display the underlying distribution from the VLQ events.

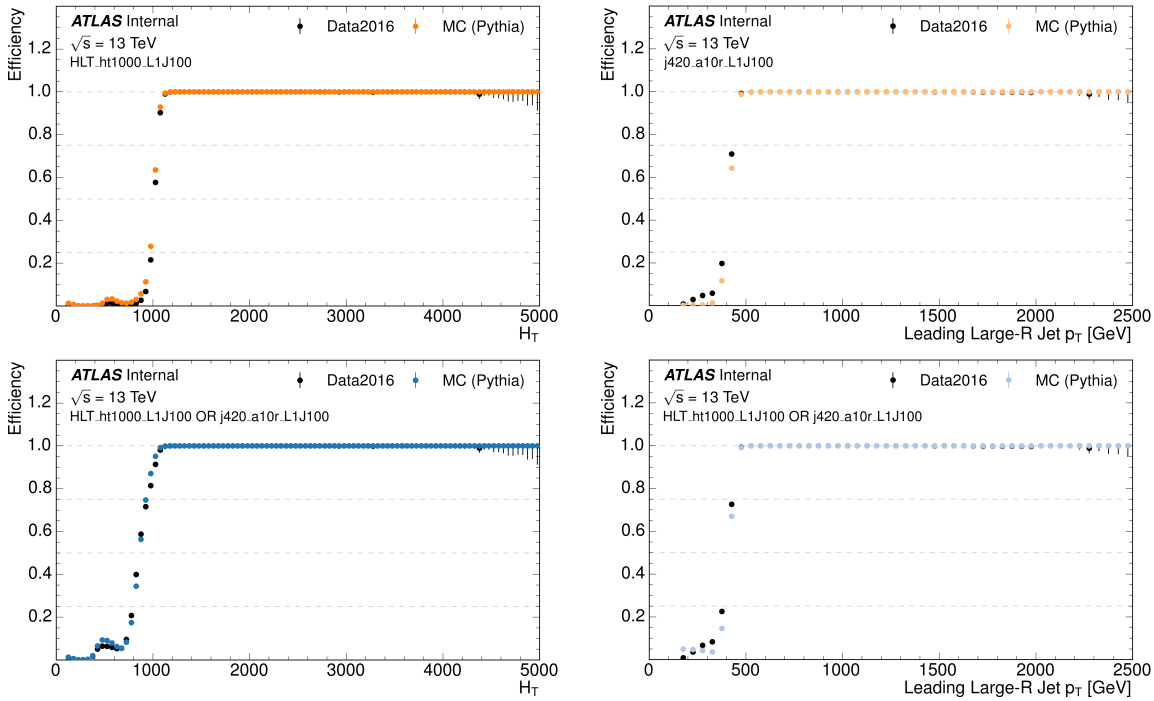

Figure 5.41: Trigger efficiencies in data and MC as a function of H_{T} (left) and leading large-R jet p_{T} (right). The top plots show the individual triggers as a function of their respective distributions and the bottom plots show the logical OR between the two triggers.

analyses, which use the 77% working point, because the multi-jet background is significantly reduced via the increase in the light jet rejection factor from 134.34 to 381.32. In addition to the standard reconstruction and calibration procedures for small-R jets, see Sect. 4.2.5, a MC-based correction on the small-R jet mass is performed. The correction reduces discrep-

ancies between geometrically-matched truth-particle jets and small-R jets in the response $m_{\text{reco}}/m_{\text{truth}}$.

A similar overlap removal algorithm as that discussed in Sect. 5.1.2 prioritizes the reconstruction of certain physics objects depending on the degree of overlap. For electron and small-R jets, the selected electron is reconstructed within $\Delta R < 0.2$ of a given selected jet, the jet is removed from the event. However, if the selected electron is reconstructed within $0.2< \Delta R <0.4$ of a selected jet, the electron is removed from the event. For muons and small-R jets, a p_{T}-dependent overlap removal working point is applied. Muons are removed if $\Delta R < \min\left[0.4, 0.04+10/p_{\text{T}}\ [\text{GeV}]\right]$ and jets are removed if one of the following conditions is met:

- The jet has < 3 tracks ($p_{\text{T}}> 500$ MeV) associated to the primary vertex or (${p_{\text{T}}}^{\mu}/\texttt{sumTrkPt500} > 0.7$ and ${p_{\text{T}}}^{\mu}/{p_{\text{T}}}^{\text{jet}} > 0.5$)
- Muon ID track ghost associated to the jet or $\Delta R(\text{jet}, \mu) < 0.2$

After the overlap removal procedure, events are required to have exactly zero leptons, ensuring orthogonality with the other VLQ analyses, and a kinematic selection on the jet p_{T} thresholds and H_{T}. The full baseline selection, after the trigger, data quality, physics object selection, and overlap removal requirements, includes

- $H_{\text{T}} > 1.15$ TeV
- Exactly zero leptons with $p_{\text{T}}>20$ GeV
- At least four small-R jets with descending minimum p_{T} thresholds of 300, 200, 125, and 75 GeV

Once the baseline selection is complete, variable–R re–clustered ('vRC') jets are constructed from the small-R jets in the event. The vRC jets incorporate two reconstruction techniques, the re–clustering and variable–R algorithms, to identify the hadronic decays of heavy objects, e.g., W, Z, or Higgs bosons and top quarks. The re–clustering algorithm [113] combines small-R jets together according to the anti-k_t algorithm to form re–clustered ('RC') jets. RC jets are different from the standard large-R jets described in Sect. 4.2.5. Rather than using topo-clusters like the standard large-R jets, RC jets cluster small-R jets, which have already been formed from topo-clusters, as the inputs to the anti-k_t clustering algorithm. A tremendous advantage to using RC jets over large-R jets concerns the treatment of systematic uncertainties. Large-R jets require their own set of calibrations and uncertainties, while RC jets directly inherit the calibrations and uncertainties from the small-R

jets. RC jets offer similar reconstruction performance as large-R jets, see Fig. 5.42, without contributing new systematic uncertainties to the likelihood fit that could reduce sensitivity to new physics [114].

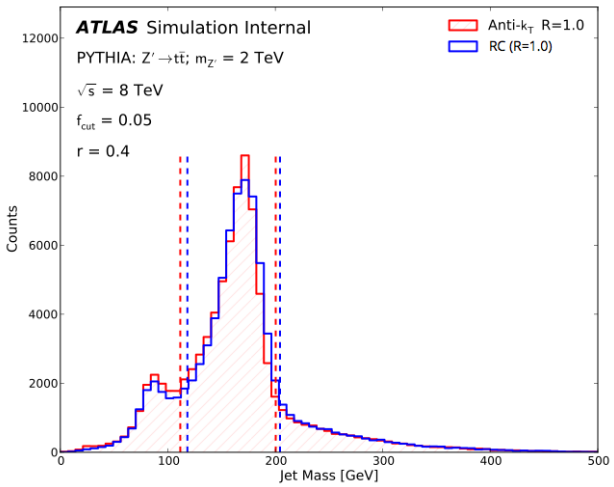

Figure 5.42: Comparison of standard large-R jets ('anti-k_t R =1.0') and RC jets ('RC (R=1.0)') using simulated events at $\sqrt{s} = 8$ TeV of the hypothetical $Z'(m = 2\text{ TeV}) \to t\bar{t}$. The vertically dashed lines depict the mass windows that contain 68% of events. RC jets and large-R jets offer similar performance, the distributions have similar shapes and structures, for reconstructing jet mass.

RC jets are common in ATLAS analyses, including recent VLQ searches [2, 3]. In the previous results, RC jets were typically used as a method to classify events using their properties, e.g., the RC jet mass. In this analysis, the re–clustering technique is expanded to include the variable–R algorithm (VR) [115] and build vRC jets. It is well-known that the separation between the decay products R of a heavy object with mass m is p_{T}-dependent: $R \sim 2m/p_{\text{T}}$ [116]. Using a fixed–radius to reconstruct RC jets can become sub-optimal at high-p_{T} or in topologies that include many reconstructed objects, e.g., VLQ decays as described in Sect. 5.1.2. Thus, the VR algorithm appropriately considers the p_{T}-dependence of boosted objects. The VR algorithm modifies the anti-k_t clustering algorithm to use a variable, rather than fixed, radius parameter. Similar to the procedure outlined in Sect. 4.2.5 and defined in Eq. 4.3, inputs to the clustering sequence are identified and their distance measures are calculated with one modification:

$$R \to R_{\text{eff}} = \frac{\rho}{p_{\text{T}}} \tag{5.14}$$

such that d_{iB} becomes $p_{T,i}^2 R_{\text{eff}}^2$ where ρ is a dimensionful constant defined as the clustering scale. The value of ρ is a free parameter of the method and an additional parameter R_{max} can be defined to regulate the maximum size of the jet, as in Eq. 5.15.

$$R_{\text{eff}} = \min\left[\frac{\rho}{p_{\text{T}}}, R_{\text{max}}\right] \tag{5.15}$$

Trimming is also applied to vRC jets using the same procedure and parameters as large-R jets, see Sect. 4.2.5, to remove sensitivity to soft radiation. However, it should be noted that the jets used to form vRC jets already have a minimum p_{T} of 25 GeV and they have passed the overlap removal and JVT cuts. Thus, trimming is effectively already applied before the re–clustering. As an example, a vRC jet with a p_{T} of at least 500 GeV would be affected by trimming if a constituent had a p_{T} of only 25 GeV.

Fig. 5.43 compares the jet mass reconstruction of vRC jets for two different values of ρ, $2m_Z$ and $2m_{\text{top}}$, with RC jets ($R = 1.0$) and large-R jets ($R = 1.0$). Both fixed–radius collections, RC jets and large-R jets, display sub-optimal performance for both leading p_{T} and sub-leading p_{T} jets in the events. The fixed–radius jets capture more energy than associated with the Z boson and top quark, as seen in the long tail of the mass distribution, while the vRC jets do not because the clustering stops once the jet energy reaches the scale ρ. The exact value of ρ can impact the reconstruction performance for vRC jets. In Fig. 5.44, vRC jets are compared with RC jets and large-R jets for different values of ρ in events where the reconstructed jets are geometrically matched to the top parton in $T\bar{T}$ events, $\Delta R(\text{top,jet}) < 0.75R$, where $R = 1.0$ for RC jets and large-R jets and $R = \rho/p_{\text{T}}$ for vRC jets. Depending on the value of ρ, the reconstructed vRC jets can significantly underestimate the top quark mass, while fixed–radius jets overestimate the mass more than vRC jets (the bottom right quadrant of each distribution in Fig. 5.44).

Given the reconstruction sensitivity to the ρ parameter in the VLQ final states, a single value of ρ was optimized for all heavy objects generated in VLQ decays (W,Z,H, and top). Distributions for different values of ρ and truth-matched jets are shown in Fig. 5.45. The best mass reconstruction was achieved by setting $\rho = 315$ GeV and $R_{\text{max}} = 1.2$ for all heavy objects. The tails in the electroweak boson jet mass distributions are less drastic for $\rho = 315$ GeV and the full top quark decay can be reconstructed.

After clustering small-R jets into vRC jets, a deep neural network (DNN) is used to classify each vRC jet into one of four categories: W/Z (V), Higgs, top, or *light*. Distinguishing between W and Z bosons is unreliable, as shown for the vRC jet mass distribution in Fig. 5.46, and they are considered as one category in the DNN. The *light* category refers to any sources of jets, quarks or gluons, produced in QCD multi-jet events.

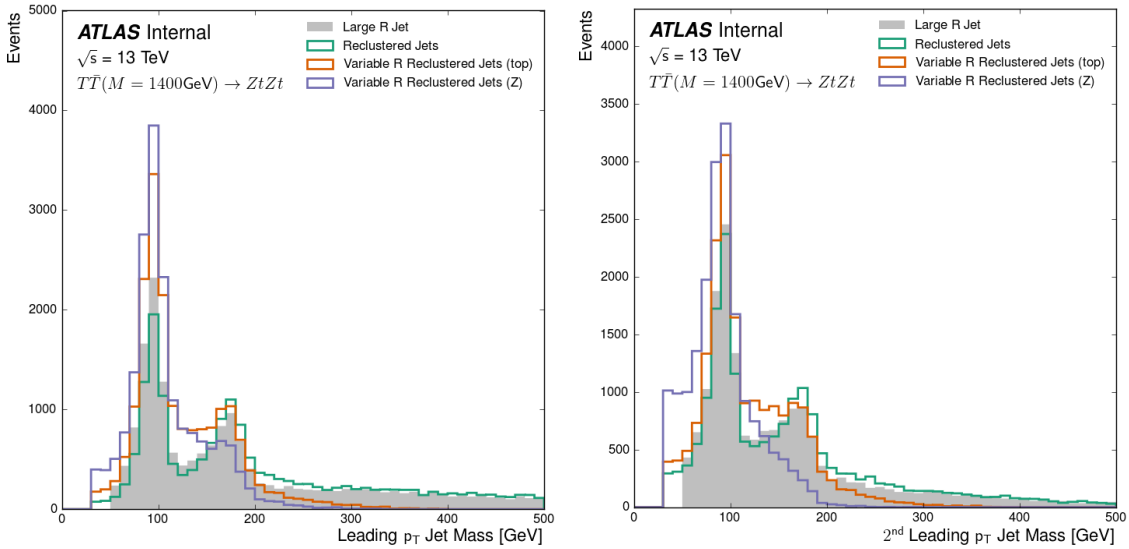

Figure 5.43: Performance of vRC jets compared to standard large-R jets and RC jets in $T\bar{T} \to ZtZt$ events ($m_T = 1.4$ TeV). The leading p_T jet mass (left) and sub-leading p_T jet mass (right) distributions demonstrate the degrading performance in a busy final state for fixed–radius (the long tails in the mass distributions) compared to variable-radius clustering.

The DNN produces output values between 0 and 1 for each class. These values are converted into a probability to improve the rejection of light jets, see Eq. 5.16. Specific signal efficiency working points have been defined for the analysis. The selections on the probability values for V, Higgs (H), and top (t) are listed in Table 5.10.

$$
\begin{aligned}
P(V) &= \log_{10}\left(\frac{\text{DNN}(V)}{0.9\cdot\text{DNN}(\text{light}) + 0.05\cdot\text{DNN}(t) + 0.05\cdot\text{DNN}(H)}\right) \\
P(H) &= \log_{10}\left(\frac{\text{DNN}(H)}{0.9\cdot\text{DNN}(\text{light}) + 0.05\cdot\text{DNN}(V) + 0.05\cdot\text{DNN}(t)}\right) \\
P(t) &= \log_{10}\left(\frac{\text{DNN}(t)}{0.9\cdot\text{DNN}(\text{light}) + 0.05\cdot\text{DNN}(H) + 0.05\cdot\text{DNN}(V)}\right)
\end{aligned}
\tag{5.16}
$$

Table 5.10: Tagging vRC jet working points for V, H, and top tagging using a multi-classification deep neural network.

Heavy Object	Threshold	Efficiency
V	$P(V) > -0.20$	70%
H	$P(H) > 0.20$	70%
top	$P(t) > -0.20$	50%

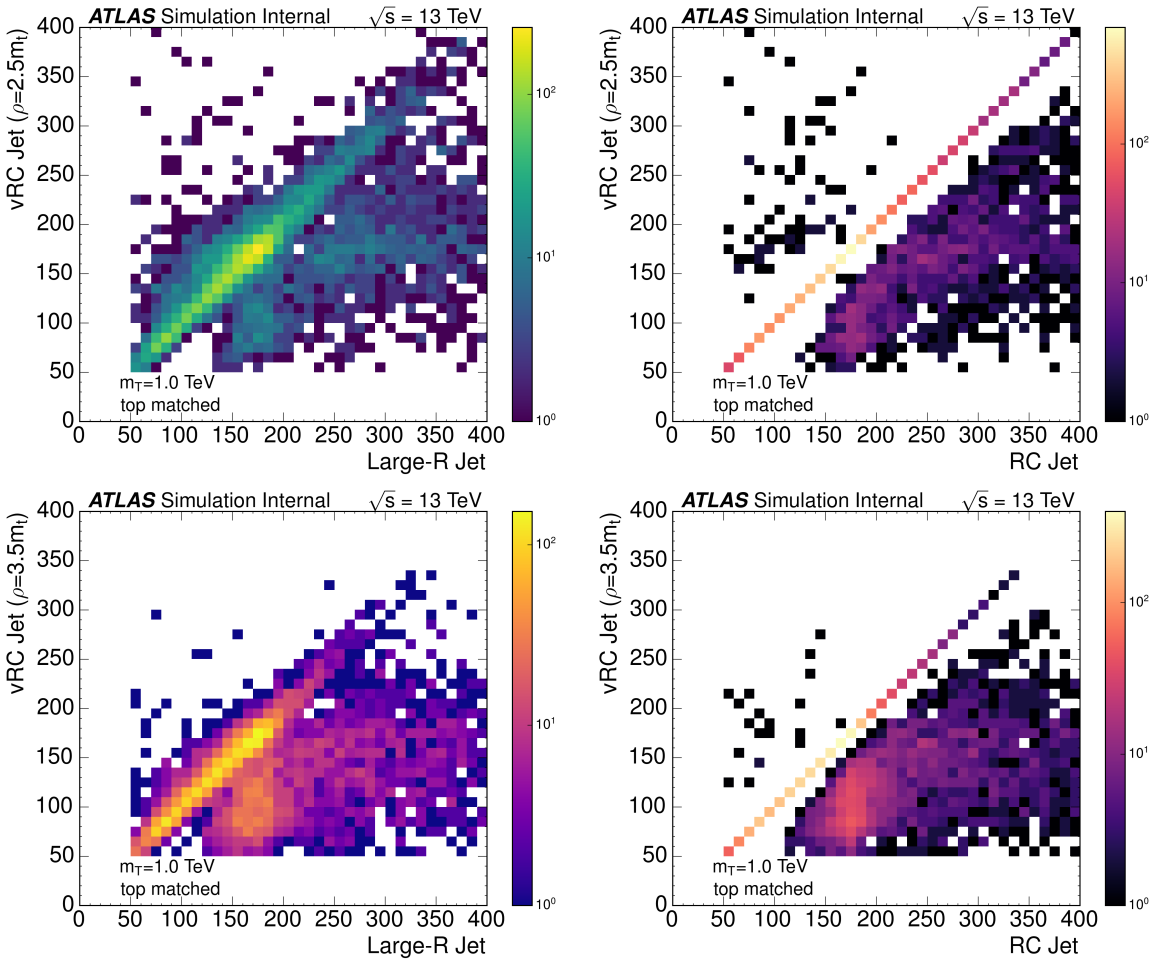

Figure 5.44: Two-dimensional histograms comparing the jet mass [GeV] for vRC jets with large-R jets (left) and RC jets (right) in $T\bar{T}$ events. The vRC jets are reconstructed with $\rho = 2.5m_{\text{top}}$ ($3.5m_{\text{top}}$) in the top (bottom) plots. Only jets that are both geometrically matched to the truth top quark are compared.

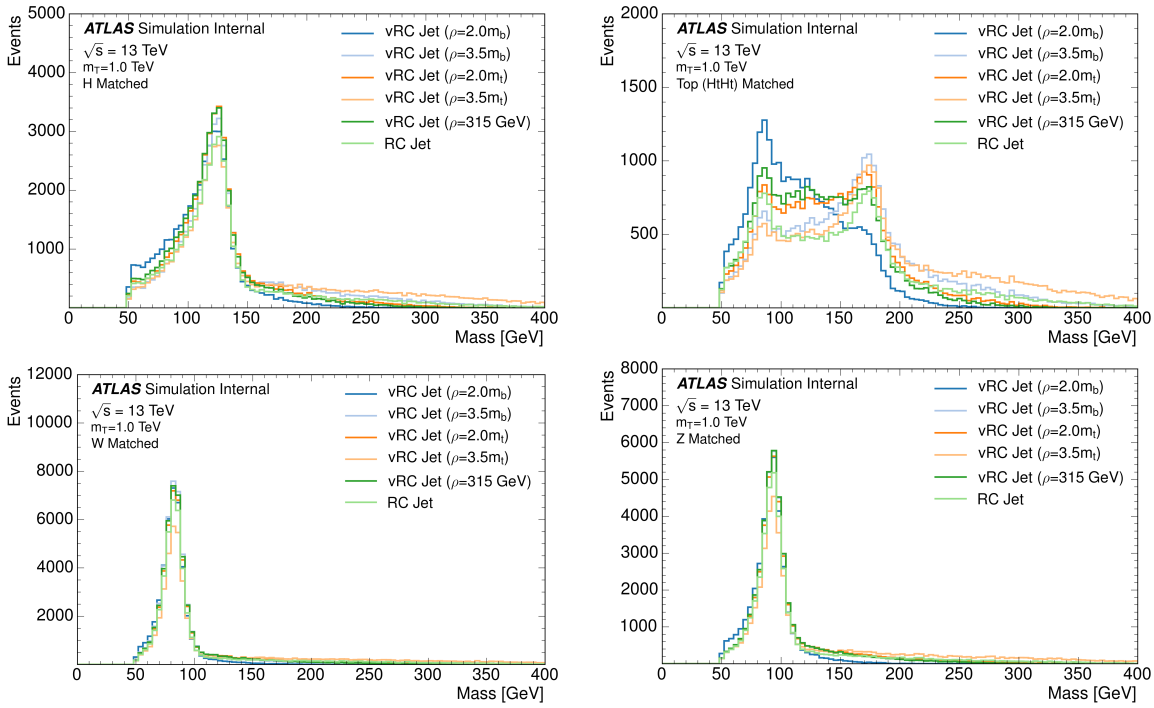

Figure 5.45: Jet mass distributions for different values of ρ (m_b =boson mass in the figure, m_t = top mass) compared to RC jets ($R = 0.8$). Each vRC jet uses $R_{\text{max}} = 1.2$ to capture lower p_{T} objects.

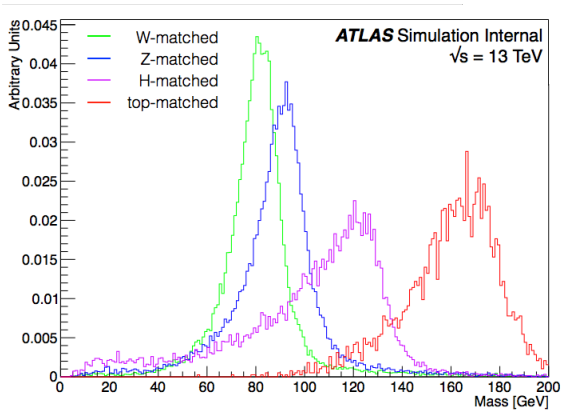

Figure 5.46: Jet mass for single small-R jets that are truth-matched to partons produced by W/Z/Higgs bosons and top quarks.

It is possible for vRC jets to be tagged as multiple objects. In those instances, an assignment to one of the tagged objects is made based on the vRC jet probability distributions, see Table 5.11. Details of the DNN architecture and inputs are described in App. A.

Table 5.11: Methodology to resolve vRC jets tagged as multiple objects.

Ambiguity	Condition	Result
V and H	$\log_{10}(\text{DNN}(V)/\text{DNN}(H)) > -0.25$	V (else: H)
V and top	–	top
H and top	$P(H) > 0.70$	H (else: t)
V,H, and top	–	H

The final step in the all-hadronic event selection requires at least two b-tagged small-R jets at the 70% efficiency working point. At least two boson–tagged vRC jets are required to improve sensitivity to the VLQ decays. Finally, $40< E_{\text{T}}^{\text{miss}} < 200$ GeV as part of the orthogonality with the high–$E_{\text{T}}^{\text{miss}}$ VLQ analysis and the data-driven multi-jet estimation.

5.2.3 Classification of Event Topologies

Events that pass the full selection are classified into 14 signal regions based on the final number of tagged bosons (VV,HH,VH), top tags (0,1,$\geq$2), and b-tags (2,$\geq$ 3). The category requiring at least two top quarks is merged for all combinations of selected bosons because this region has a low number of events when split by boson type that causes instability in the data-driven background estimate. This classification enhances the significance relative to the SM background and, in the event of a discovery, the VLQ decays are separated into channels that closely reflect the branching ratios. In the following sections, the regions are named to reflect the bosons, tops, and b-tags, e.g., "VV0t2b" represents the regions with 2 V tags, 0 top tags, and 2 b-tags. The 2 top tag region uses 'X' to represent all bosons W, Z, and Higgs, not to be confused with the VLQ X with electric charge +5/3.

Control regions in this analysis are used to validate the data driven background estimation described in Sect. 5.2.4. Performance of the data-driven technique using the 1 b-tag control region is shown in Fig. 5.47. Non-closure of the data-driven estimate in the 1 b-tag control region, i.e., the disagreement between data and prediction, is propagated as an uncertainty to the 2 and 3 b-tag signal regions, described further in Sect. 5.2.5.

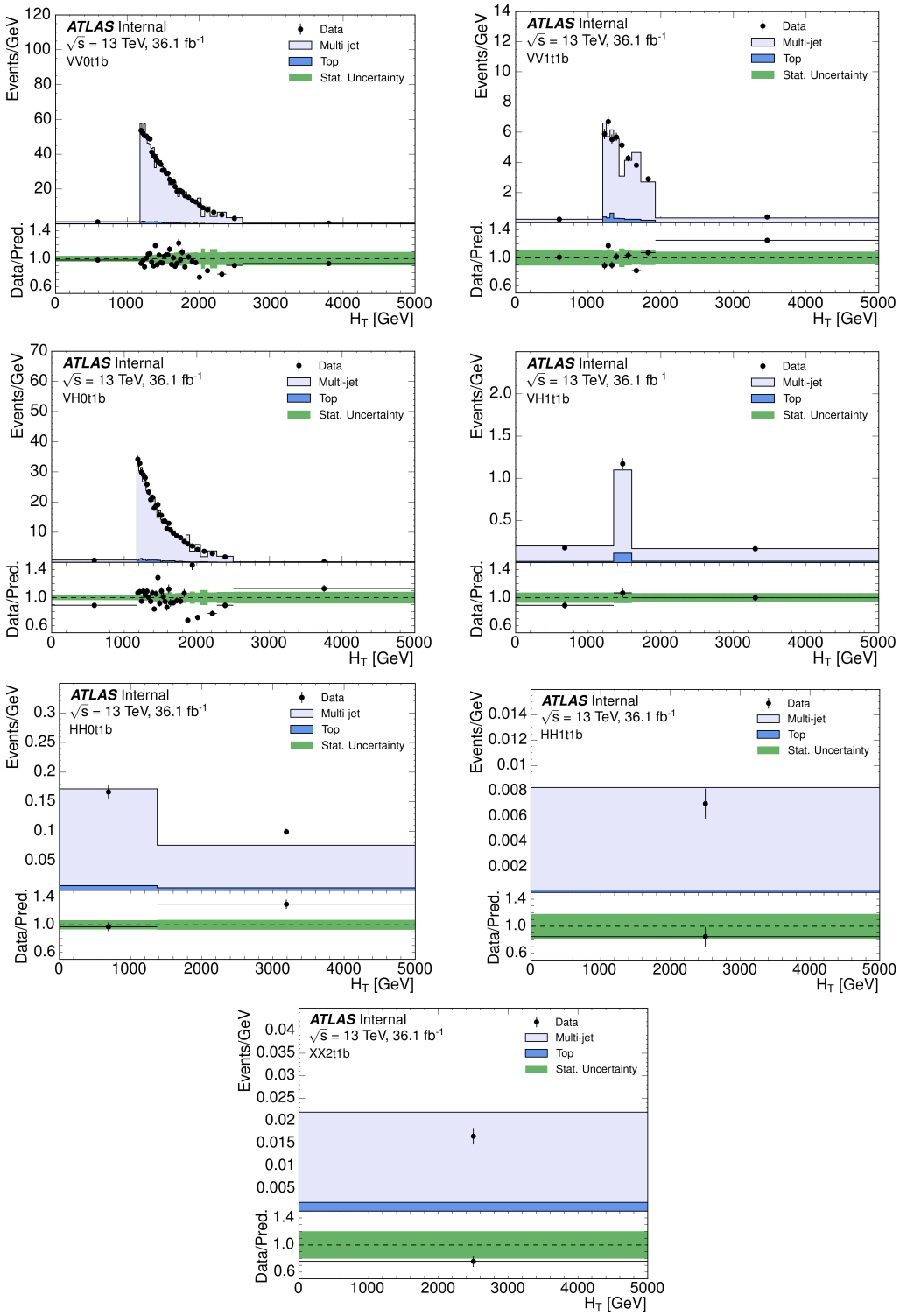

Figure 5.47: Distributions in H_{T} comparing Data and prediction for events in the 1 b-tag control regions separated by number of boson and top tags. Each bin is normalized to the bin width. Only statistical uncertainties are shown in green.

The signal acceptance (fraction of total events that pass the full selection) is shown in Figs. 5.49–5.55. The signal acceptances are displayed for each analysis region and VLQ decay channel, for both $T\bar{T}$ and $B\bar{B}$ events. Across all regions, the $B\bar{B} \to HbHb$ and $T\bar{T} \to HtHt$ are the decay modes with the largest acceptance, as shown in Fig. 5.48.

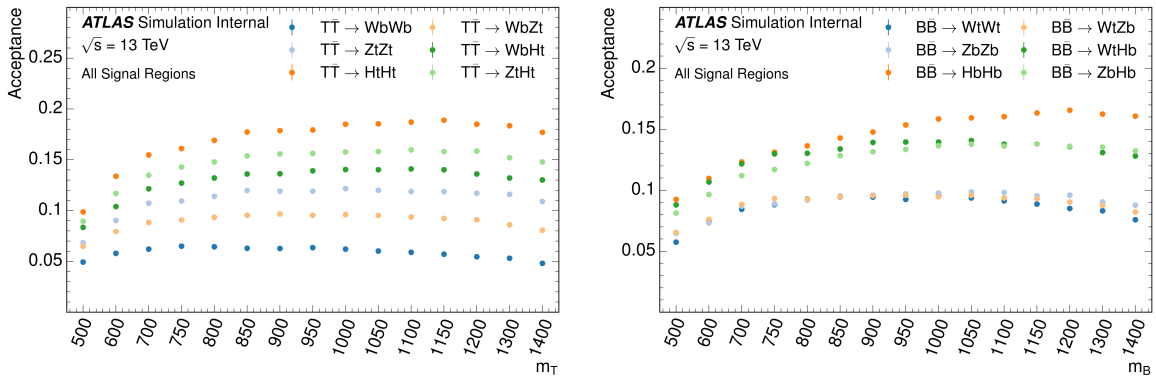

Figure 5.48: Total signal acceptance for all $T\bar{T}$ (left) and $B\bar{B}$ (right) decays across all signal regions. The VLQ decays to Higgs bosons yields the largest signal acceptance.

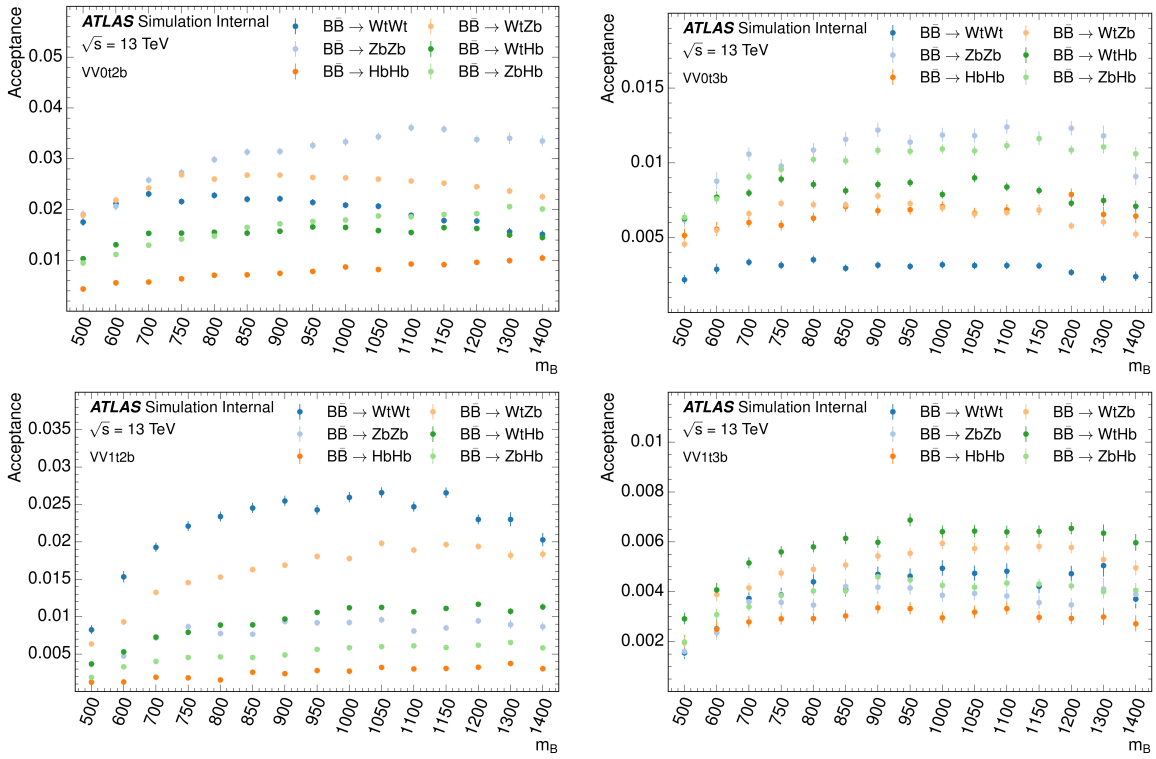

Figure 5.49: Signal acceptance for all $B\bar{B}$ decays in the VV regions: 0 top-tags and 2 b-tags (upper left), 0 top-tags and ≥ 3 b-tags (upper right), 1 top-tag and 2 b-tags (lower left), 1 top-tag and ≥ 3 b-tags (lower right).

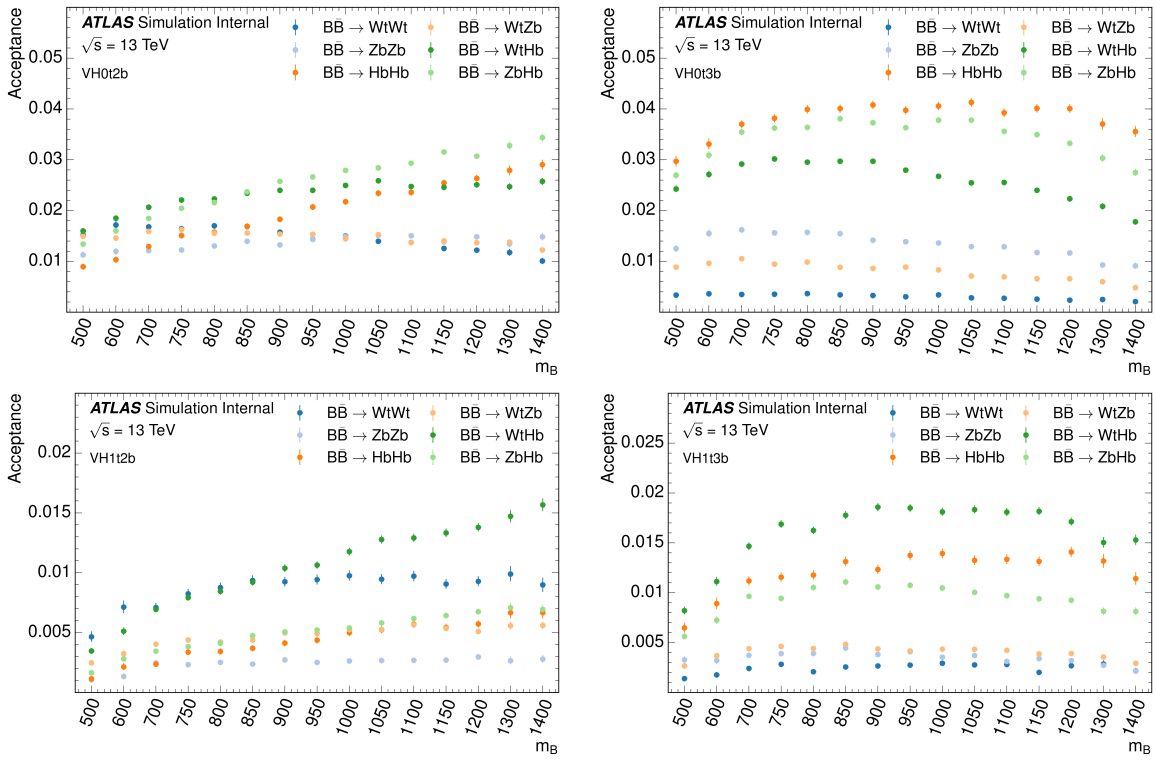

Figure 5.50: Signal acceptance for all $B\bar{B}$ decays in the VH regions: 0 top-tags and 2 b-tags (upper left), 0 top-tags and ≥ 3 b-tags (upper right), 1 top-tag and 2 b-tags (lower left), 1 top-tag and ≥ 3 b-tags (lower right).

5.2.4 Background Estimation

A mixture of simulated and data-driven techniques are used to study the SM background contributions to signal regions. The different processes and their respective estimation techniques are described below.

5.2.4.1 Simulation of SM backgrounds

SM backgrounds from $t\bar{t}$, single top, and $t\bar{t} + X$ are studied with simulated events, described in Sect. 5.2.1. The normalization and shape are both taken directly from simulation for all of these processes.

5.2.4.2 Data-driven background estimation

A data-driven technique is necessary to accurately predict the normalization and shape of the expected SM multi-jet background. The multi-jet background is estimated separately for each of the fourteen signal regions using the boson tagging quality and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 'ABCD' method. The regions of the ABCD method are defined by the boson tagging quality and

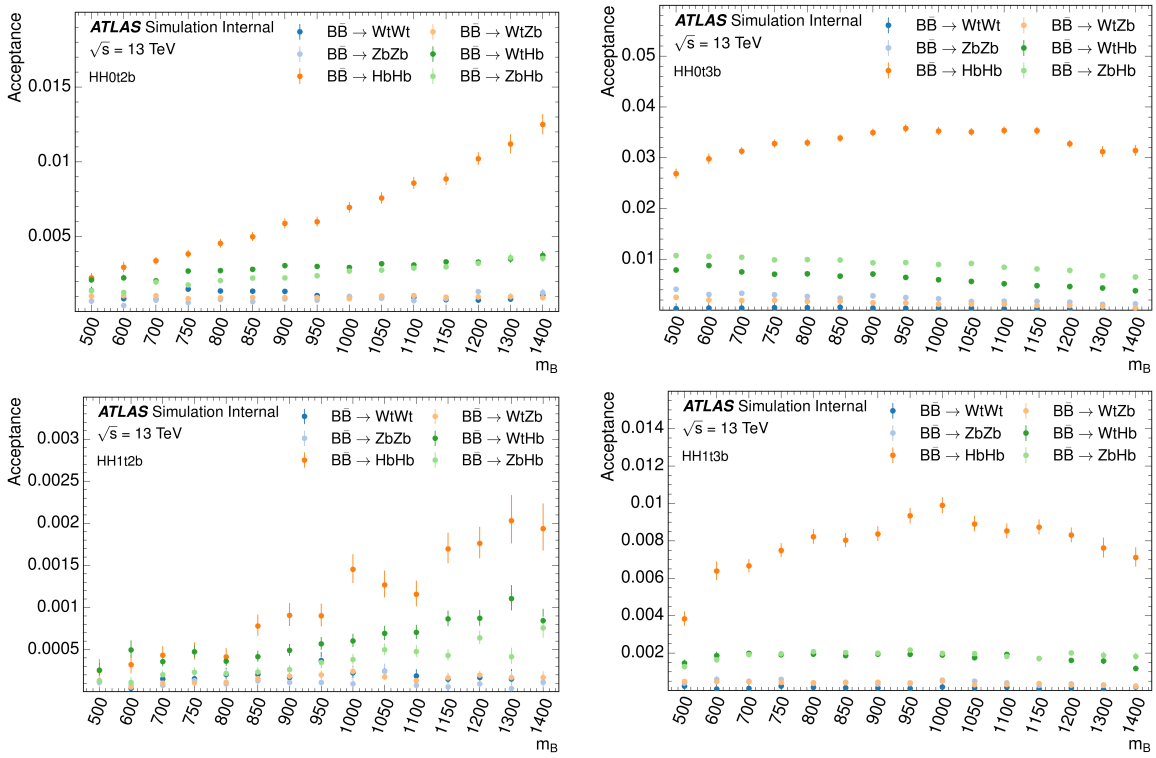

Figure 5.51: Signal acceptance for all $B\bar{B}$ decays in the HH regions: 0 top-tags and 2 b-tags (upper left), 0 top-tags and ≥ 3 b-tags (upper right), 1 top-tag and 2 b-tags (lower left), 1 top-tag and ≥ 3 b-tags (lower right).

$E_{\mathrm{T}}^{\mathrm{miss}}$, diagrammed in Fig. 5.56. The boson tagging quality axis uses two boson tagging categories: *loose* and *tight*. The tight tagging definition was defined in Sect. 5.2.2, and the loose tagging is defined using the vRC jet mass, m_j:

- V Boson: $69 < m_j < 104$ GeV
- Higgs Boson: $104 < m_j < 155$ GeV

The mass windows are derived from the jet mass for single small-R jets that are geometrically matched to all partons from VLQ decays[4] and contain 68% of events, see Fig. 5.46. The tagging window for the Higgs boson is slightly less than 68% due to the orthogonality requirement, i.e., one jet cannot be tagged as both a V and Higgs boson. The two boson tagging quality regions are

1. Regions B & D: 2 tight tags
2. Regions A & C: <2 tight tags `AND` 2 loose+tight tags

[4]To be geometrically matched, the partons p_i are required to be within $\Delta R(p_i, \text{jet}) < 0.3$ of the small-R jet.

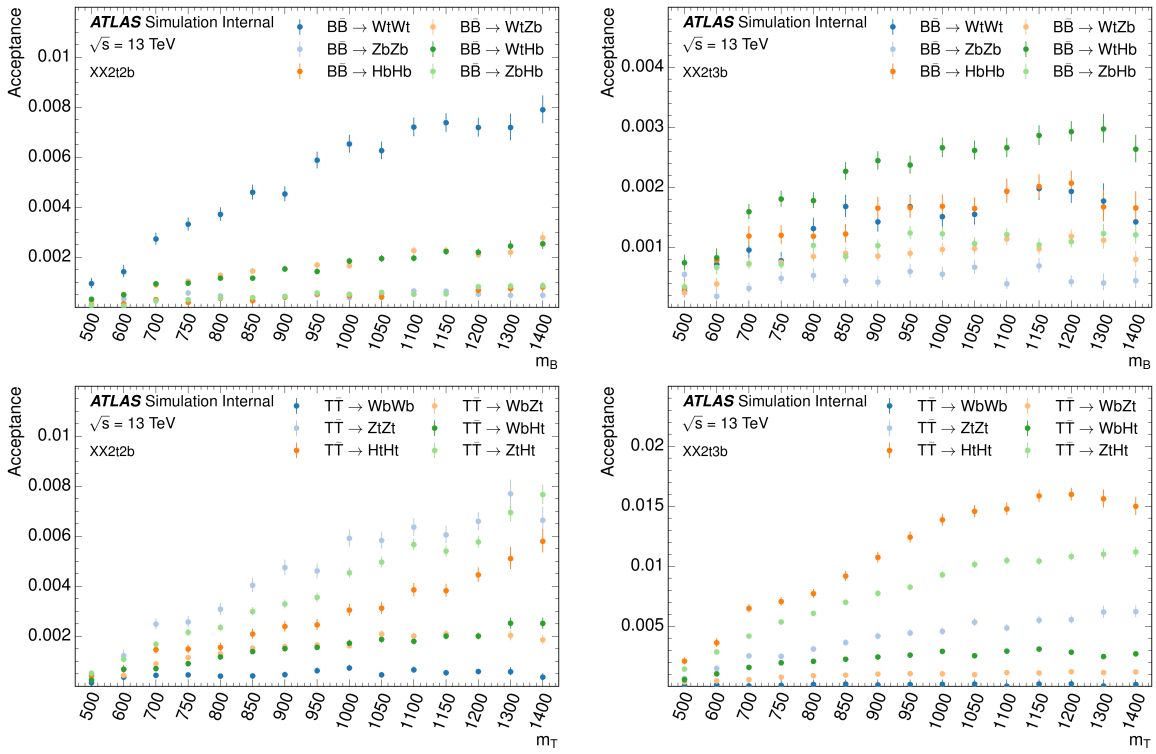

Figure 5.52: Signal acceptance in the two top tag regions: 2 b-tags (upper left) and ≥ 3 b-tags (upper right). The $B\bar{B}$ events are drawn in the top two figures and the $T\bar{T}$ events are drawn in the bottom two figures.

Regions A and C are allowed to have at most one tight boson tag because it improves the statistics used in the multi-jet estimation and reduces differences in the kinematics between the regions. The second axis, $E_{\mathrm{T}}^{\mathrm{miss}}$, is uncorrelated with the boson tagging axis. A boundary at $E_{\mathrm{T}}^{\mathrm{miss}} = 40$ GeV separates control and signal regions and $E_{\mathrm{T}}^{\mathrm{miss}}$ has an upper bound of 200 GeV due to the orthogonality requirement with the high–$E_{\mathrm{T}}^{\mathrm{miss}}$ analysis. Simulated multi-jet events have a correlation factor of –0.05 between these two variables across all analysis regions. The distribution for multi-jet events in the ABCD regions are drawn in Fig. 5.56. Semi-leptonic events, with a soft or mis-reconstructed lepton, and $Z \to \nu\nu$ decays are sources of real $E_{\mathrm{T}}^{\mathrm{miss}}$ in all-hadronic VLQ events. SM multi-jet events are not expected to have any significant amounts of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the slight differences between the VLQ and expected multi-jet shapes are exploited for the multi-jet background estimation. Fig. 5.56 diagrams the specific regions used to estimate both the normalization and shape of the multi-jet background.

In regions A, B, and C the multi-jet contribution is estimated by subtracting all simulated MC events from the measured data. The relationship between the signal region, D, and the control regions is given by $D = C \times (B/A)$. A binned estimation gives the shape of H_{T}

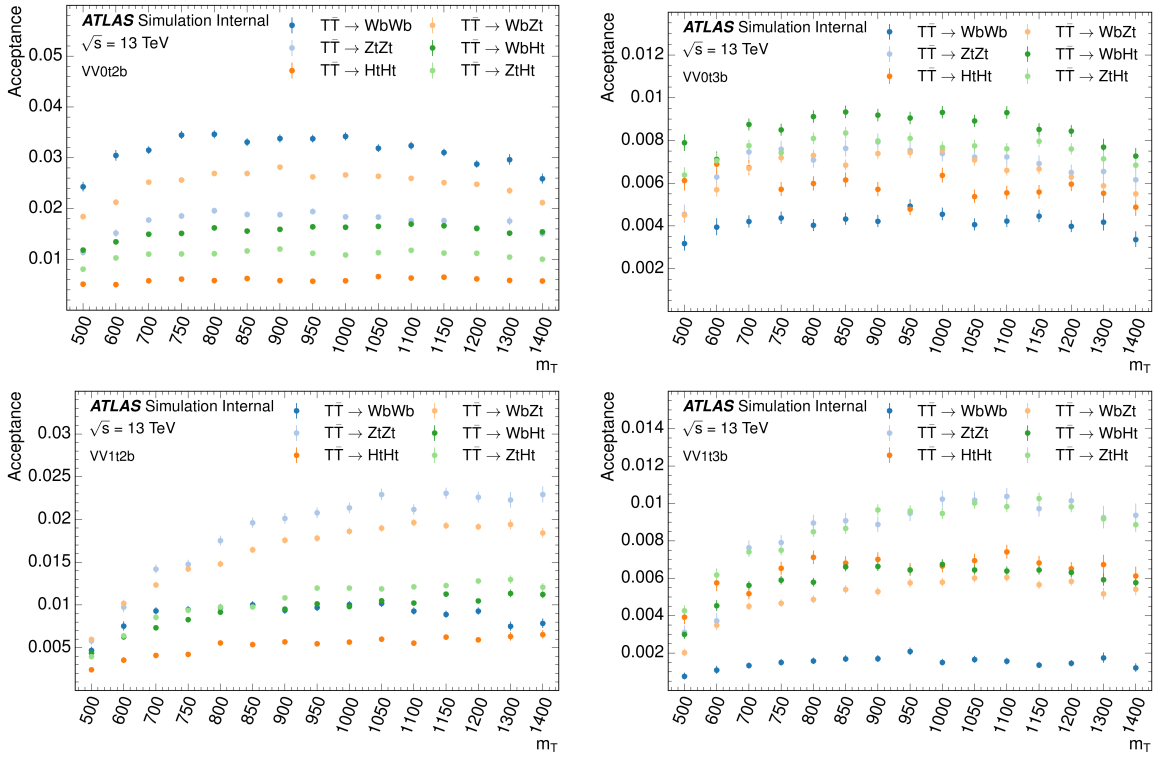

Figure 5.53: Signal acceptance for all $T\bar{T}$ decays in the VV regions: 0 top-tags and 2 b-tags (upper left), 0 top-tags and ≥ 3 b-tags (upper right), 1 top-tag and 2 b-tags (lower left), 1 top-tag and ≥ 3 b-tags (lower right).

in region D, and the total yield is calculated from the integral over the full distribution. Uncertainties in the multi-jet estimation are described in Sect. 5.2.5.

5.2.5 Systematic Uncertainties

This section describes the sources of systematic uncertainty considered in this analysis. These uncertainties can be broken down into several broad categories: luminosity and cross section uncertainties, detector-related experimental uncertainties, modeling uncertainties on simulated background processes, and uncertainties on data-driven background estimations. Each source of uncertainty is treated as a nuisance parameter in the fit of H_{T}.

5.2.5.1 Luminosity and cross section uncertainties

The uncertainty in the total combined 2015+2016 integrated luminosity is 3.2%. It is derived, following a methodology similar to that detailed in [101], from a preliminary calibration of the luminosity scale using $x - y$ beam-separation scans performed in August

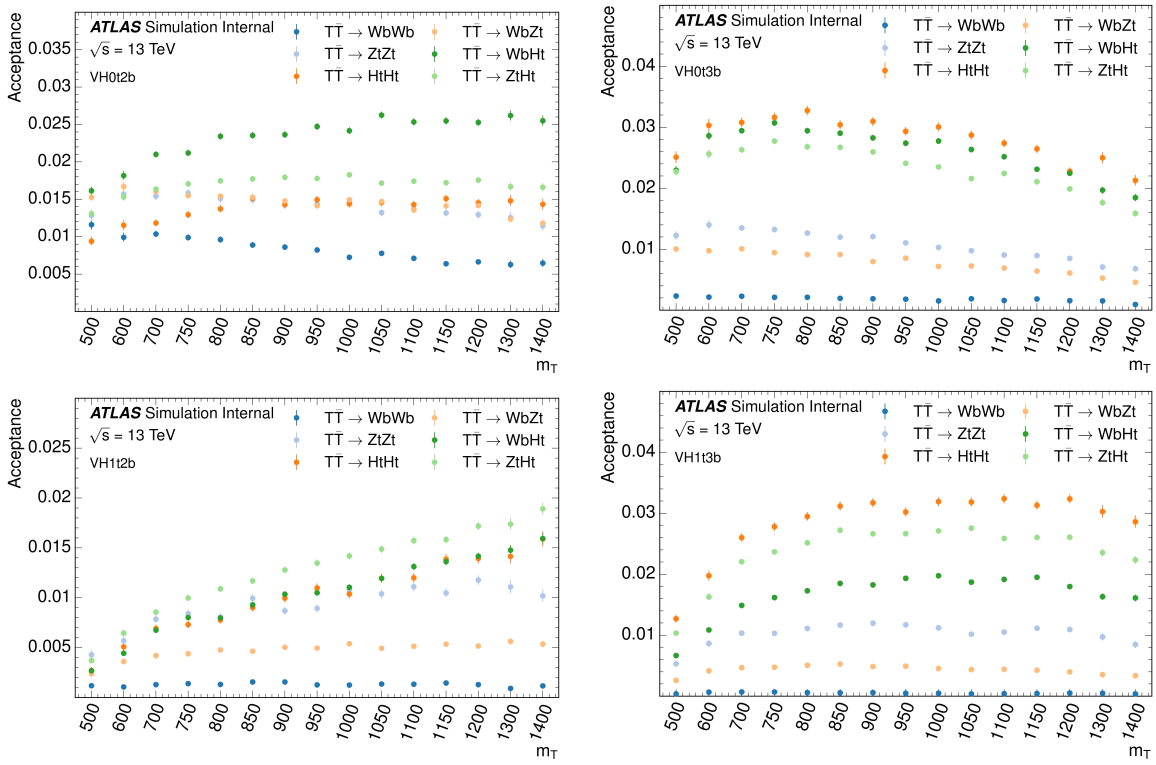

Figure 5.54: Signal acceptance for all $T\bar{T}$ decays in the VH regions: 0 top-tags and 2 b-tags (upper left), 0 top-tags and ≥ 3 b-tags (upper right), 1 top-tag and 2 b-tags (lower left), 1 top-tag and ≥ 3 b-tags (lower right).

2015 and May 2016[5]. This systematic uncertainty is applied to all backgrounds that are estimated using simulated samples and normalized to the measured integrated luminosity.

Theoretical cross section uncertainties have been used for the simulated samples that are in used in this analysis. The uncertainty on $t\bar{t}$ and single top production is 6% [44, 104, 105]. There is no consideration for the uncertainties on the $t\bar{t}$ modeling or heavy-flavor composition.

5.2.5.2 Detector-related uncertainties

Many of the detector-related uncertainties in this analysis are identical to those considered in the one lepton analysis. The description of the uncertainties for electrons, muons, $E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jet energy scale and resolution, b-tagging, and pile-up are left to Sect. 5.1.5.2 and the unique uncertainties are discussed here.

This analysis does not include large-R jets, and the uncertainties are not considered. The vRC jets do not contribute new uncertainties to the analysis as their uncertainties are

[5]This paragraph is directly copied from the ATLAS luminosity for physics definition, as recommended by the ATLAS publications committee: `https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics`

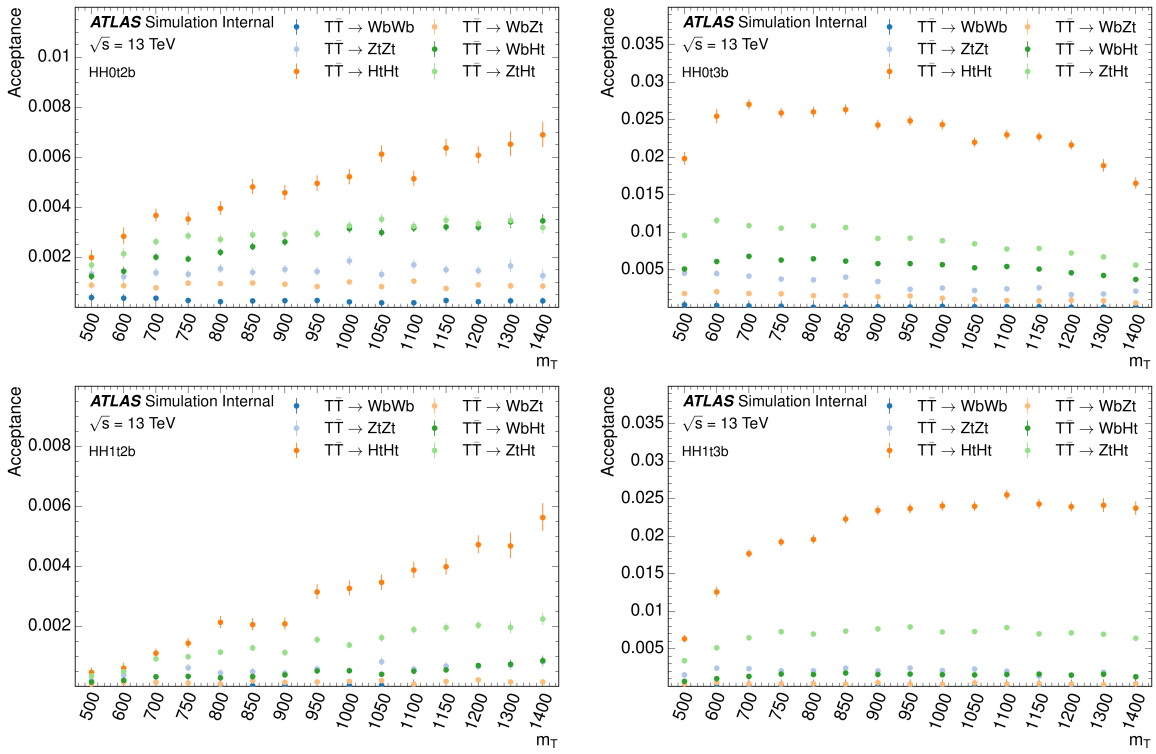

Figure 5.55: Signal acceptance for all $T\bar{T}$ decays in the HH regions: 0 top-tags and 2 b-tags (upper left), 0 top-tags and ≥ 3 b-tags (upper right), 1 top-tag and 2 b-tags (lower left), 1 top-tag and ≥ 3 b-tags (lower right).

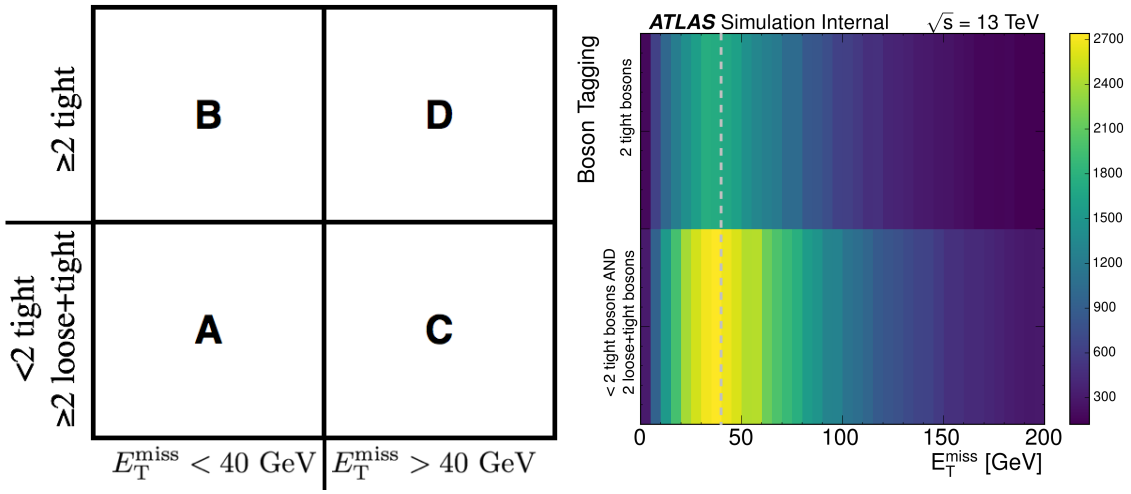

Figure 5.56: Regions definitions for the ABCD estimation of the multi-jet background (left) and the distribution of simulated multi-jet events in the ABCD regions (right). Events are binned according to the boson tagging and $E_{\mathrm{T}}^{\mathrm{miss}}$. The right plot is drawn with a gray dashed line to designate the $E_{\mathrm{T}}^{\mathrm{miss}}$ boundary between regions.

derived from the small-R jets. There are, however, additional uncertainties for the small-R jet

mass estimated due to the corresponding calibration, described in Sect. 5.2.2. Four separate components address the sources of uncertainty from the jet mass calibration: the baseline uncertainty comparing the difference between data and prediction, comparison of different MC generators, uncertainties from reconstructed tracks, and the statistical uncertainty on the results. Due to low statistics in the estimation of the mass uncertainties, the value of the uncertainty for small-R jets with p_{T}=100 GeV is extrapolated to a 50% uncertainty at p_{T}=20 GeV. In this analysis, only the fourth leading-p_{T} jet and lower are sensitive to this extrapolation. The groups of uncertainties with the largest impact are those associated with jet energy and mass scale and resolutions and b-tagging.

5.2.5.3 Uncertainties on Data-Driven Methods

The dominant multi-jet background is estimated using a data-driven technique, see Sect. 5.2.4. To quantify an uncertainty on this method, the difference in 1 b-tag control region, shown in Sect. 5.2.3, between the prediction and data is propagated to the signal regions. This uncertainty, referred to later as 'Multi-jet Closure', is the dominant uncertainty in the analysis. Fig. 5.47 contains the control region distributions and the discrepancy between data and prediction drawn in the sub-plot is directly applied to the signal regions with the relevant binning. Uncertainties on the simulated events are also propagated through the multi-jet estimation to test the sensitivity of the ABCD method to the detector-level uncertainties.

5.2.6 Statistical Analysis

This analysis uses the same methodology and tools described in Sect. 5.1.6 to test the VLQ signal hypothesis. Thus, the relevant sections that have already been described in detail are only briefly summarized.

A simultaneous profile likelihood fit is performed in all fourteen signal regions, defined in Sect. 5.2.3, using the H_{T} distribution to test for the presence of signal. Binning the H_{T} distribution is optimized separately for each region to account for different amounts of statistics, with a requirement of at least 100 events per bin. To reduce the number of nuisance parameters and improve the stability in the fit, the $t\bar{t}$ (all-hadronic and non-all-hadronic), single top, and $t\bar{t}+X$ samples are merged into one category 'Top'.

The systematic uncertainties and the corresponding nuisance parameters used in the fit are listed in section 5.2.5. Systematics that impact the normalization or shape of the prediction by less than 1% are pruned before the fit is performed. All one-sided uncertainties, those that do not have both an 'up' and 'down' variation, are subjected to smoothing and

symmetrization. In addition, all double-sided systematics are forced to be symmetrized, but no further smoothing algorithm is applied.

5.2.7 Results

Before unblinding the data in the signal regions, expected limits were derived using the total expected background as pseudo-data. Deriving expected limits on $T\bar{T}$ and $B\bar{B}$ production is critical to understand the behavior and impact of nuisance parameters in the likelihood fit. In addition, the performance of the dijet background across the different signal regions can be validated. After unblinding the data in the signal region, a simultaneous fit to all signal regions is performed to derive limits on $T\bar{T}$ and $B\bar{B}$ production.

5.2.7.1 Fit to Asimov data

Similar to the procedure described in the one lepton analysis in Sect. 5.1.7, a likelihood fit is performed using Asimov data under the background-only hypothesis. The results of the Asimov fit for the fully-hadronic final state are shown in Figs. 5.57–5.63. Fitting to Asimov data offers a test of the systematic uncertainties, pulls and constraints, and to understand the expected sensitivity to a signal hypothesis. The $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) = 1) signal hypothesis is not drawn on the post-fit plots due to the background-only fit. In each distribution, the combined backgrounds of $t\bar{t}$, single top, and $t\bar{t} + X$ are labeled as 'Top'.

Uncertainties below the pruning threshold are omitted from the fit and uncertainties can also be symmetrized or smoothed, as described in Sect. 5.1.6.1. The summary of this pruning step is illustrated in Fig. 5.64. The fitted nuisance parameters are shown with their pulls and constraints compared to the nominal value of 0 and standard deviation of 1 in Fig. 5.65. A fitted error of 1 indicates that the data has not enough statistical power to reduce the original uncertainty. As a further cross check of the fit results, the correlations between the nuisance parameters in the Asimov fit are shown in Fig. 5.66. Only nuisance parameters with a correlation coefficient of at least 10% with any other parameter are displayed. The correlations are measured by their impact on a product over all fitted bins. A ranking of the nuisance parameters according to the absolute post-fit effect on the fitted signal strength $\hat{\mu}$ in the fit to Asimov data for the signal-plus background hypothesis, described in Sect. 5.1.7.1, can be seen in Fig. 5.67, where the signal hypothesis is $B\bar{B}$ ($m_B = 1$ TeV; BR($T \to Hb$) = 1).

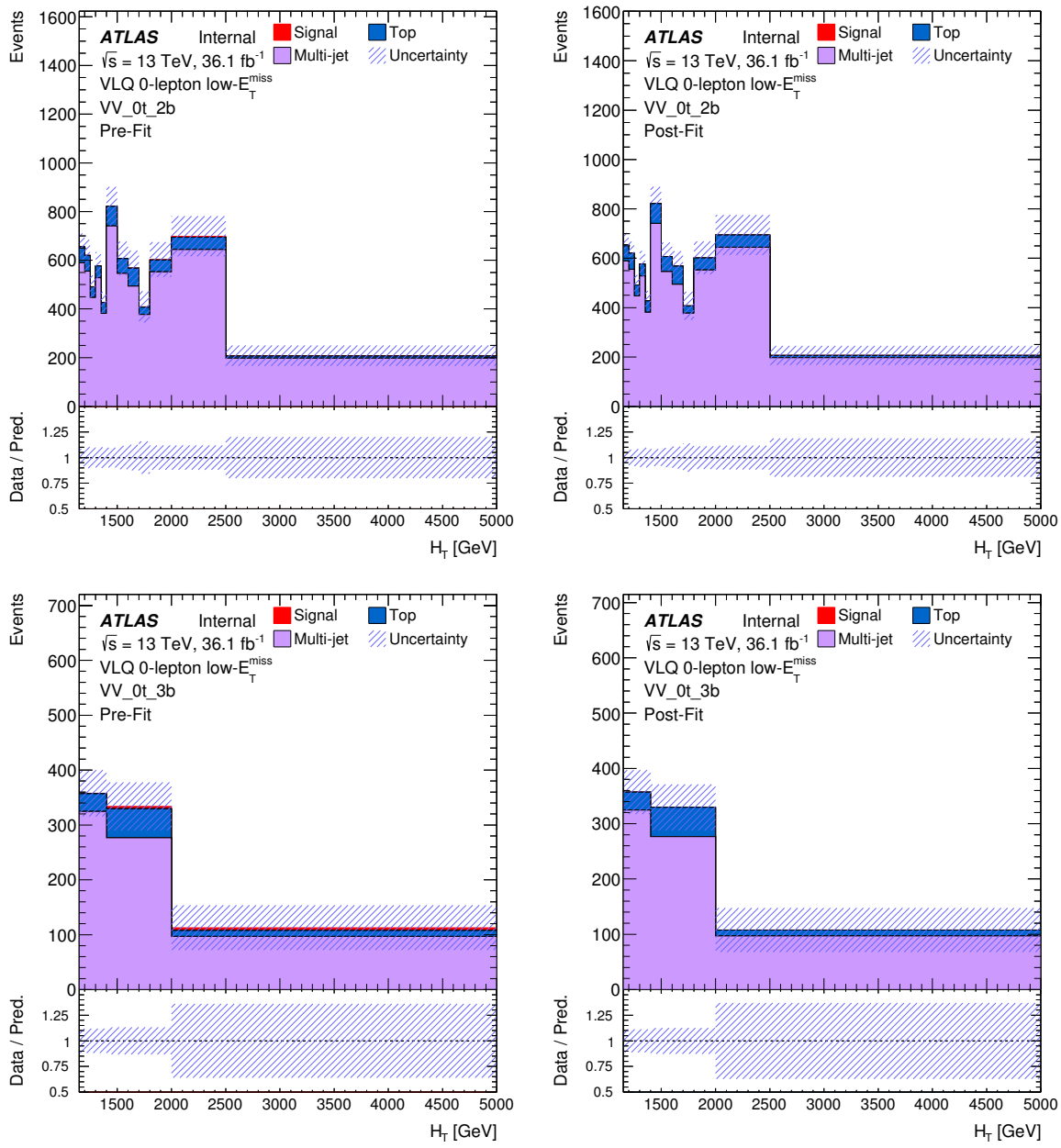

Figure 5.57: Prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the VV signal regions with 0 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

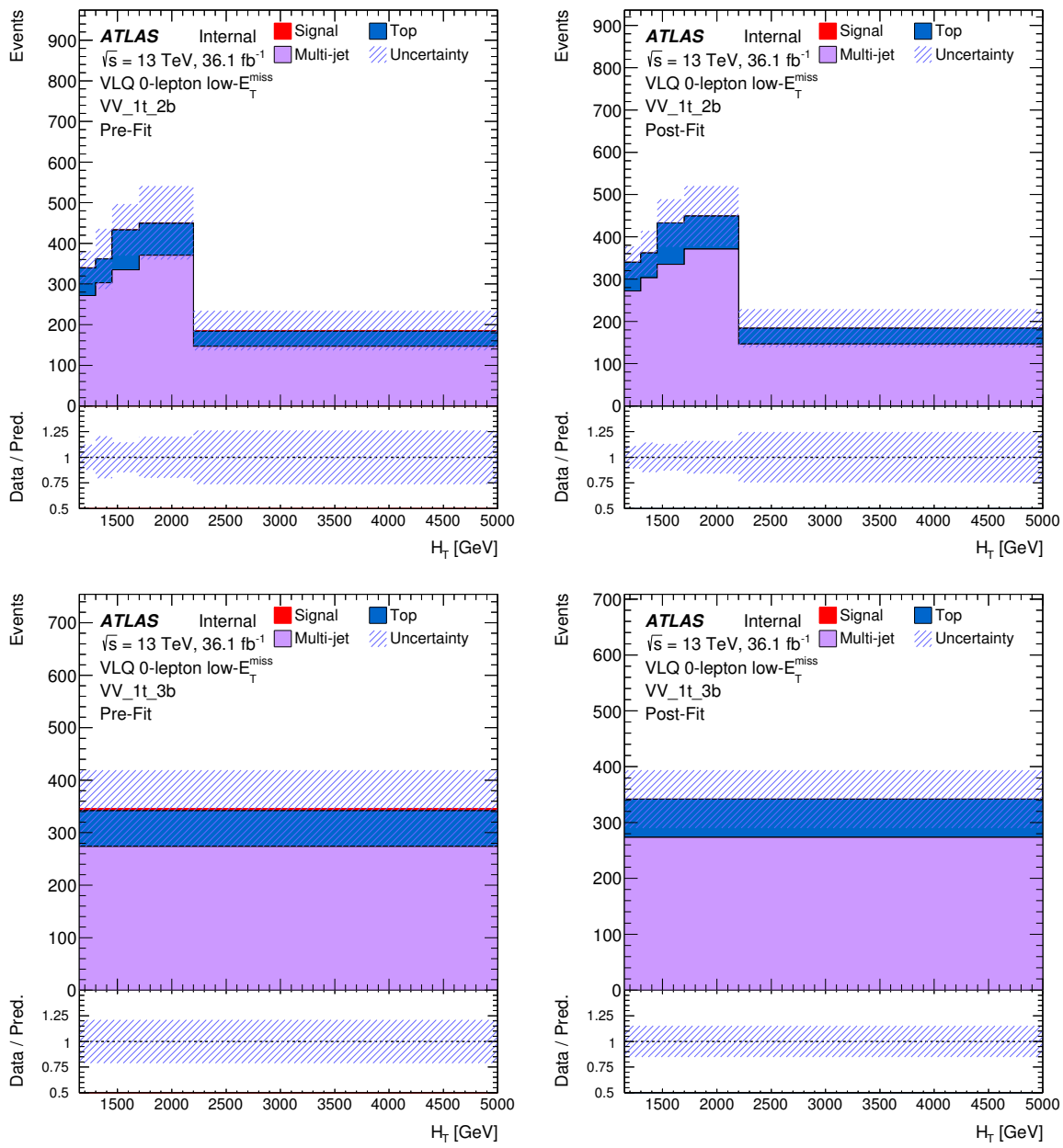

Figure 5.58: Prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the VV signal regions with 1 top tag: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

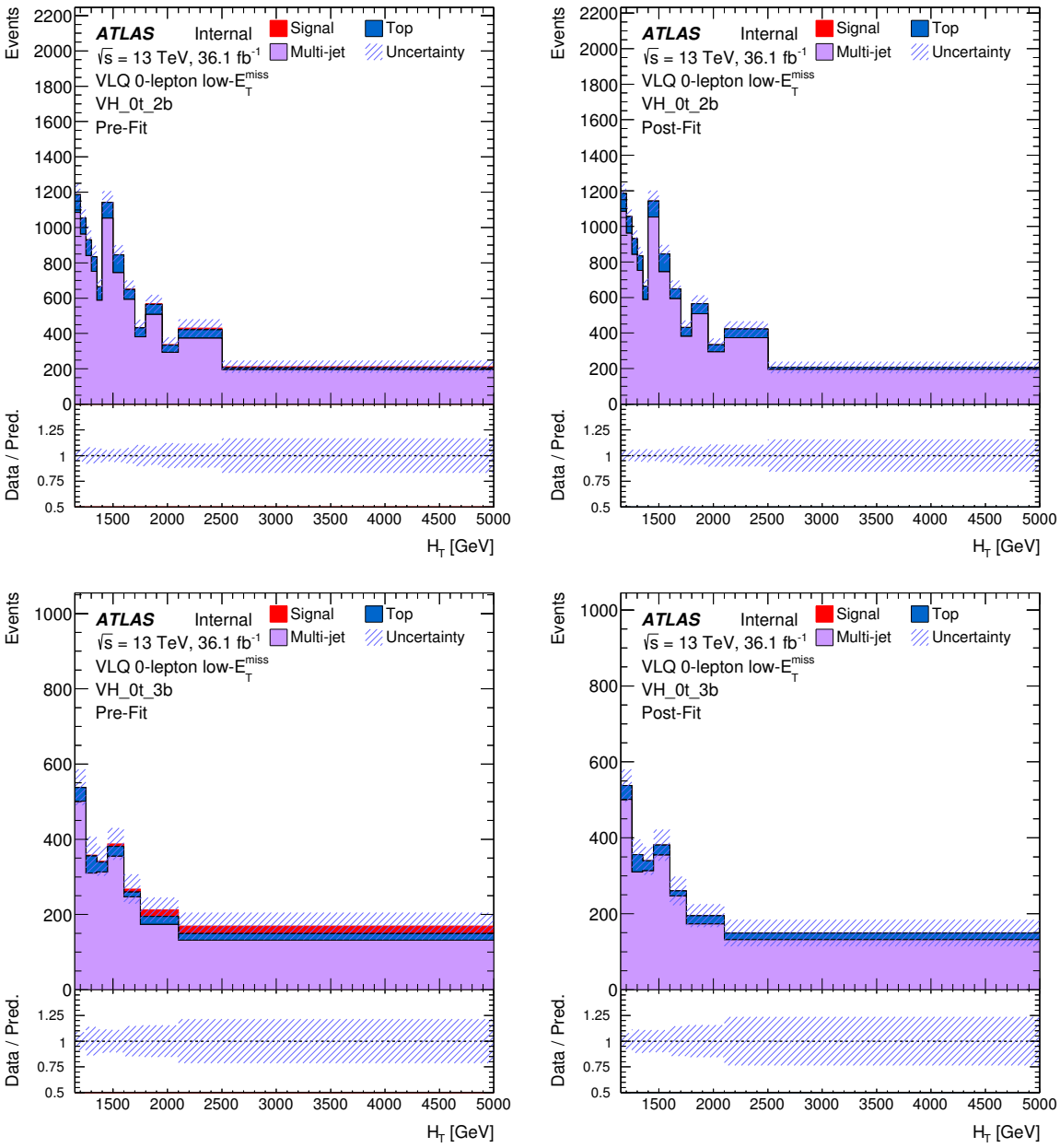

Figure 5.59: Prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the VH signal regions with 0 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

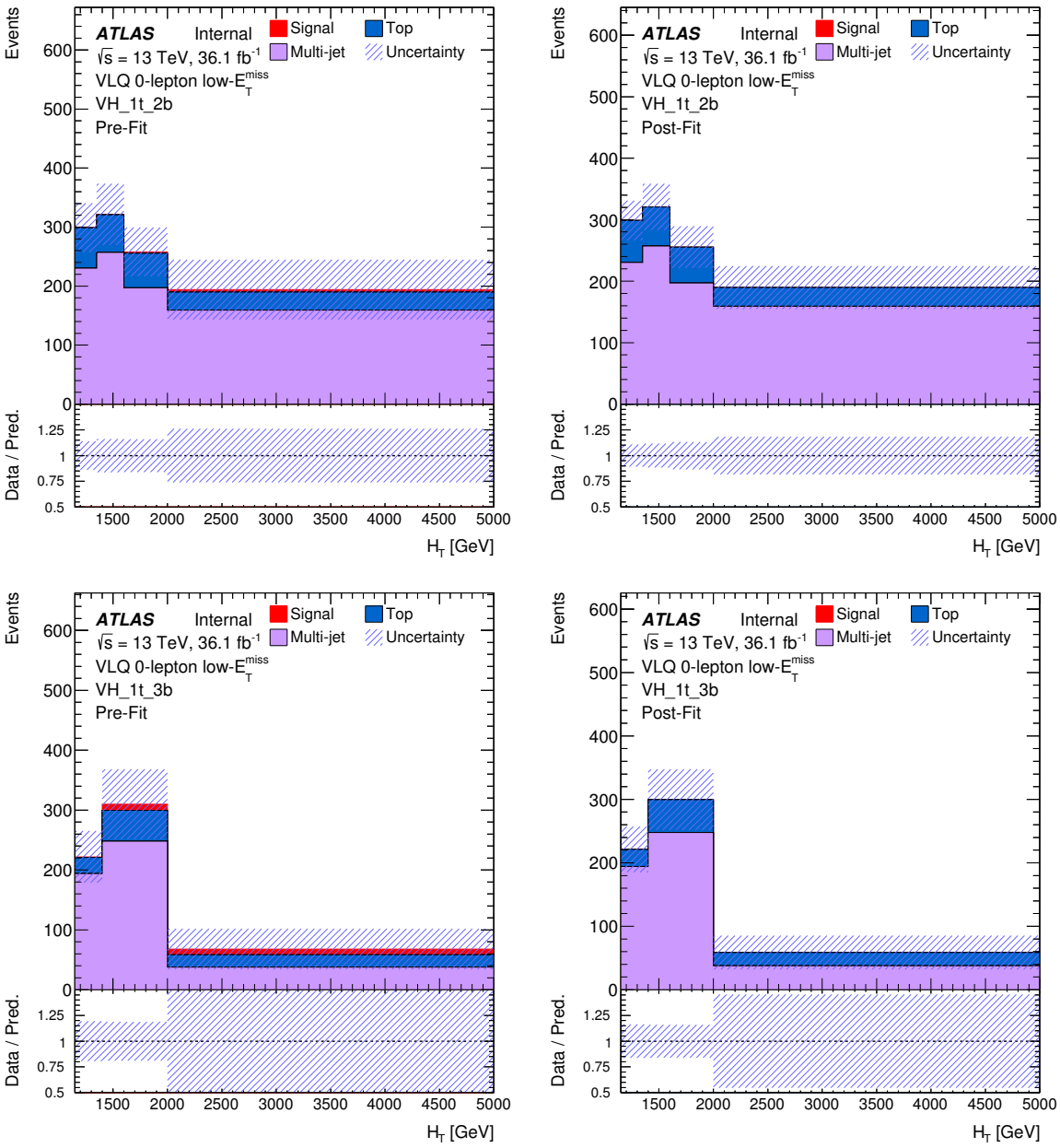

Figure 5.60: Prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the VH signal regions with 1 top tag: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

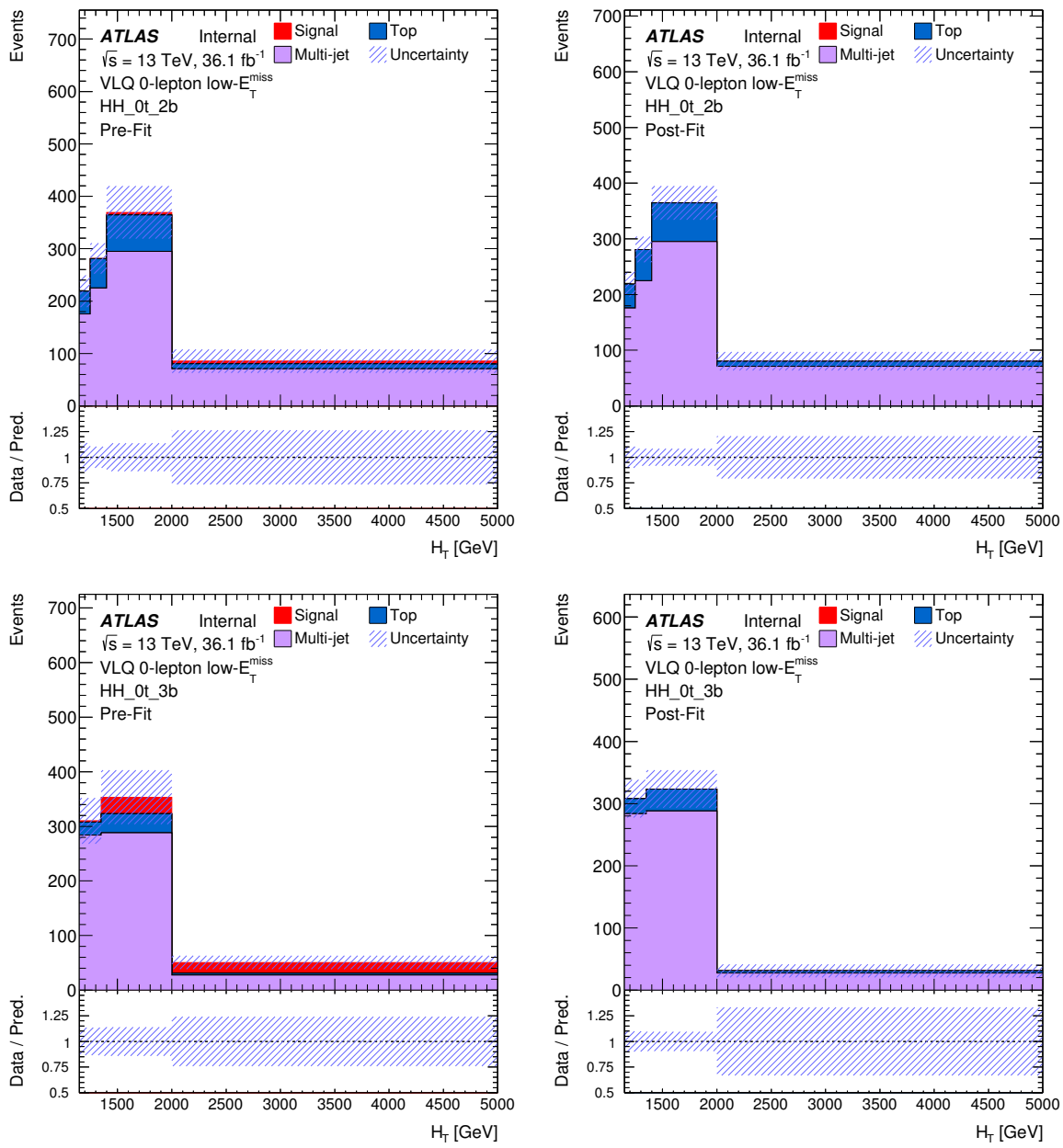

Figure 5.61: Prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the HH signal regions with 0 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

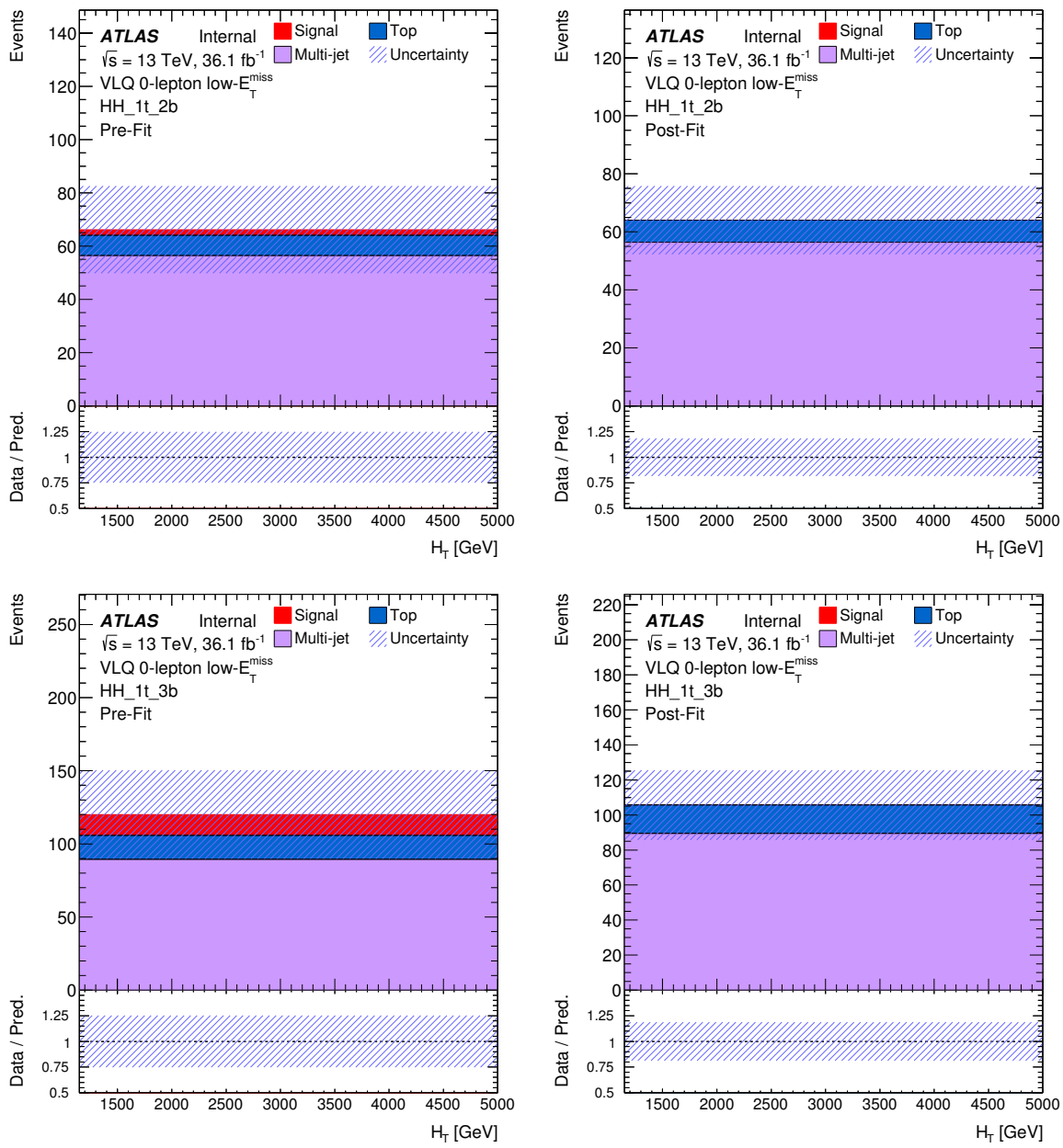

Figure 5.62: Prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the HH signal regions with 1 top tag: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

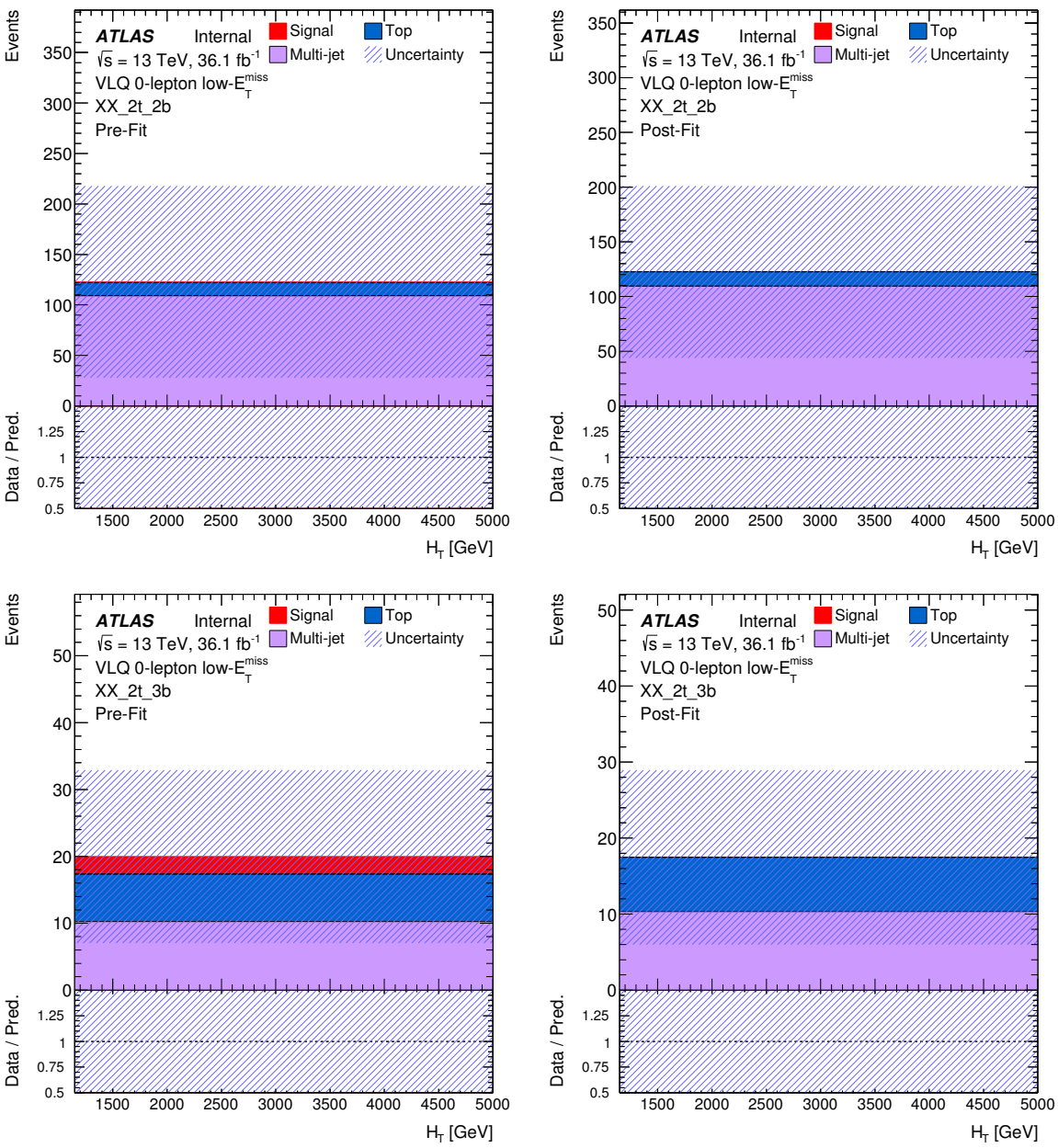

Figure 5.63: Prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the signal regions with ≥ 2 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

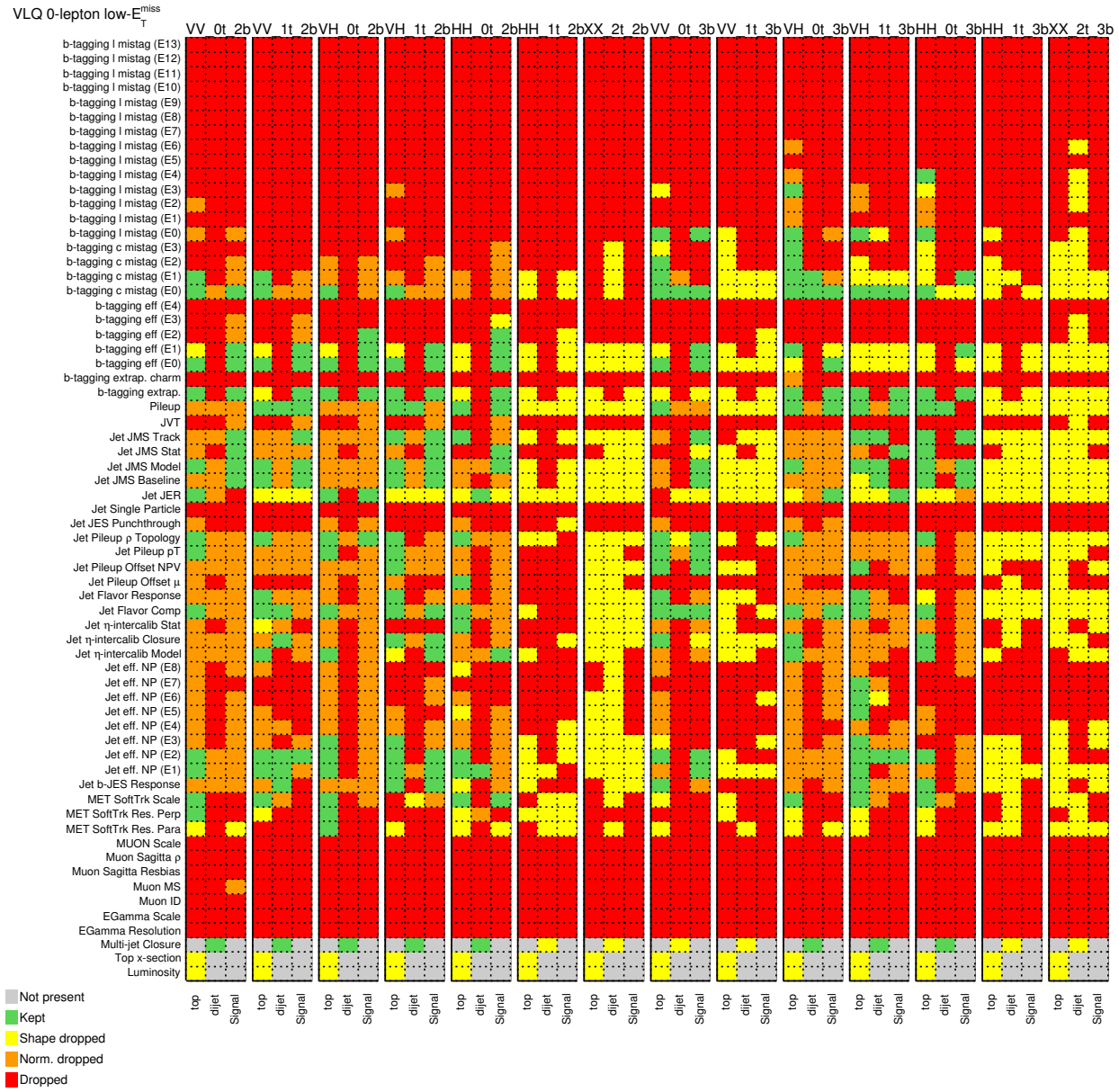

Figure 5.64: Summary of the pruning of the systematic uncertainties. The uncertainties that have the shape (normalization) component removed are yellow (orange) and those that are red were completely dropped from the fit. Uncertainties are green if both shape and normalization remain in the fit.

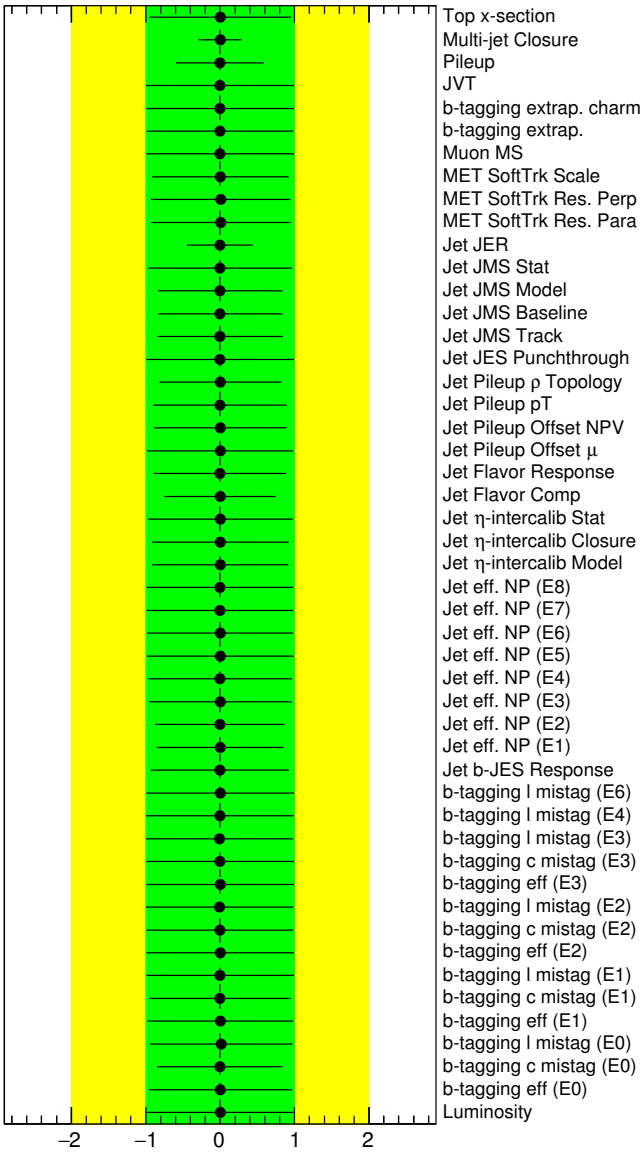

Figure 5.65: Fitted nuisance parameters in the background-only hypothesis to the Asimov dataset. As can be expected, all nuisance parameters fit at the expectation of 0 for the Asimov dataset. The constraints of the nuisance parameters can be read directly from the figure.

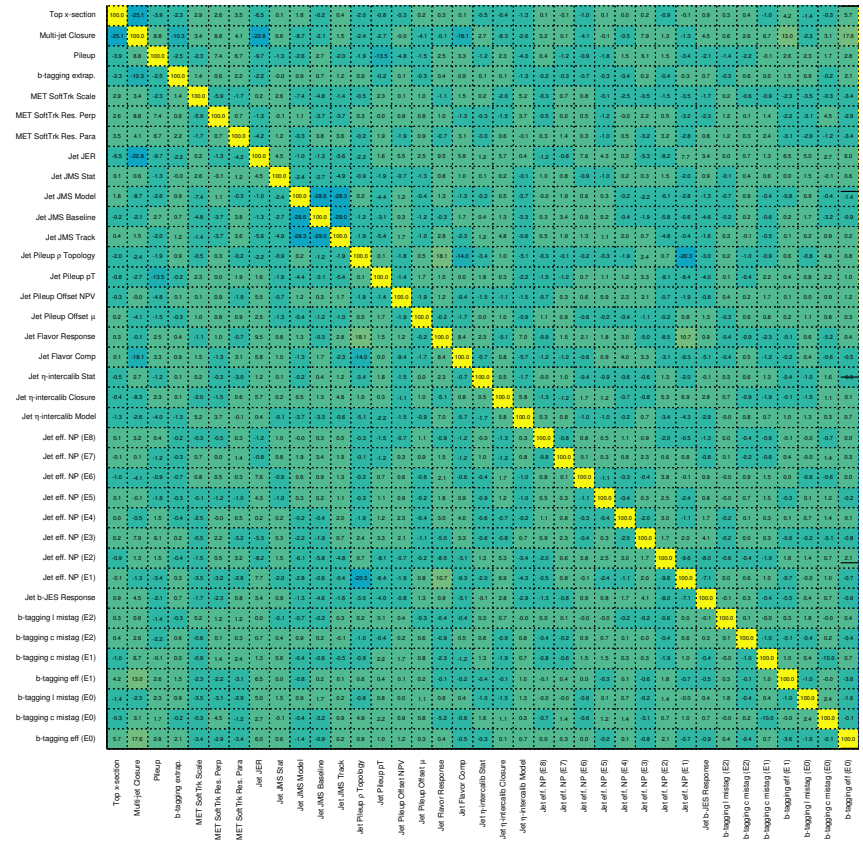

Figure 5.66: Correlation matrix showing the correlations of the nuisance parameters for the fit to the Asimov dataset in the signal+background hypothesis. The correlation scale legend corresponds to 100% (yellow), 0% (green) and -100% (blue). Only nuisance parameters with a correlation coefficient of at least 10% with any other parameter are displayed.

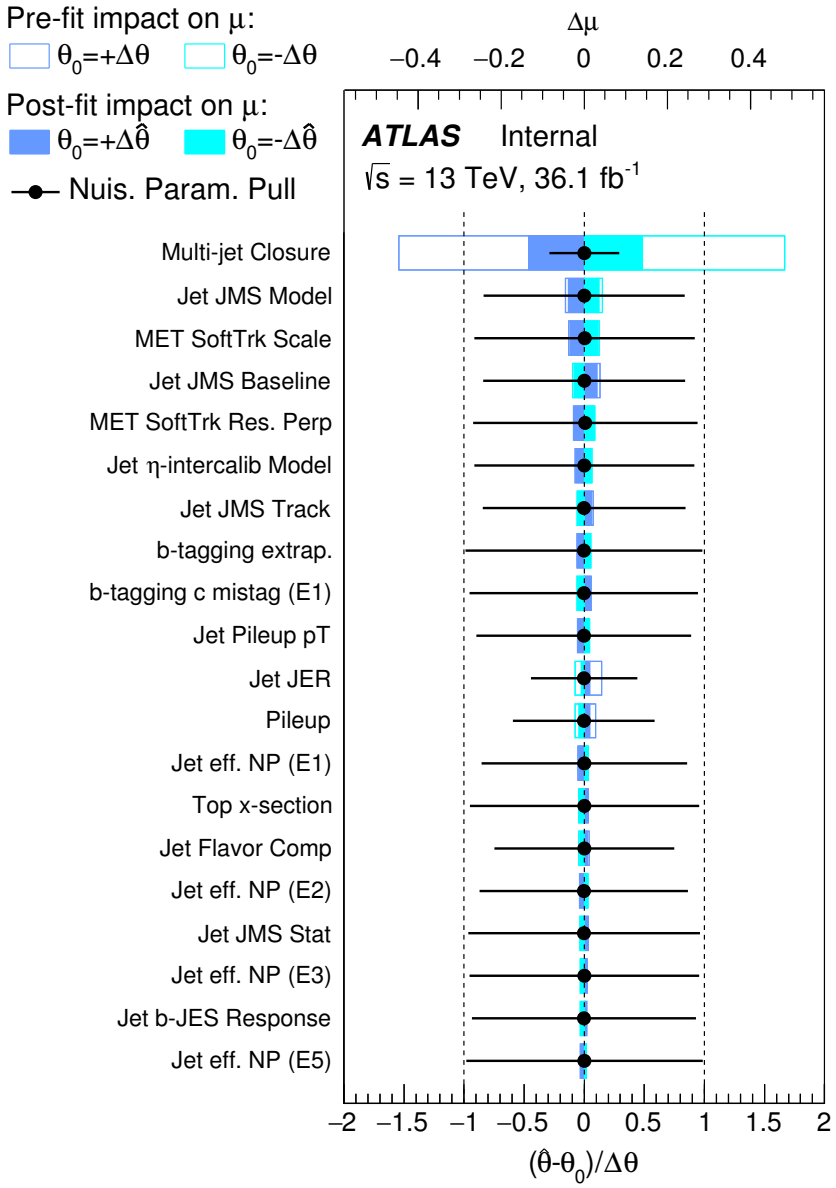

Figure 5.67: Ranking of nuisance parameters based on the fit to Asimov data in the signal-plus background hypothesis according to their effect on the uncertainty on μ ($\Delta\mu$).

Upper limits at the 95% confidence level (CL) on the $T\bar{T}$ and $B\bar{B}$ production cross section are set for three benchmark scenarios as a function of T (B) VLQ mass m_T (m_B) and compared to the theoretical prediction from `top++` [44], see Figs. 5.77 and 5.78. The resulting lower limit on m_T and m_B corresponds to the central value of the theoretical cross-section prediction. The scenarios considered involve different assumptions on the decay branching ratios. To probe the complete branching ratio plane spanned by both processes, the signal samples are re-weighted by the ratio of the respective branching ratio to the original branching ratio in PROTOS. Then, the complete analysis is repeated for each grid point. Fig. 5.79 shows the corresponding expected mass limits on T (B) in the BR($T(B) \to Ht(Hb)$) versus BR($T(B) \to Wb(Wt)$) plane, obtained by linear interpolation of the calculated CL_s versus m_T (m_B).

5.2.7.2 Fit to Unblinded Data

After unblinding the data, the same analysis performed for Asimov data is repeated with real data. A comparison of the H_{T} distributions in each signal region, both pre- and post-fit to data are shown in Figs. 5.68–5.74 and summarized in Table 5.12.

Upper limits at the 95% confidence level (CL) on the $T\bar{T}$ and $B\bar{B}$ production cross section are set for two benchmark scenarios as a function of T (B) VLQ mass m_T (m_B) and compared to the theoretical prediction from `top++` in Figs. 5.77 and 5.78. The resulting lower limit on m_T and m_B corresponds to the central value of the theoretical cross-section prediction. Table 5.13 shows the observed and expected 95% CL limits obtained for T and B quarks under different branching ratio assumptions. Due to an underestimation of the data in multiple signal regions, as shown in the post-fit plots above, the observed limits are lower than the expected limits.

Table 5.12: Event yields in all fourteen signal regions after the fit to the background-only hypothesis. The uncertainties include statistical and systematic uncertainties. The uncertainties on the individual background components can be larger than the uncertainty on the sum of the backgrounds due to correlations.

Region	Top	Multi-jet	Total background	Data
VV0t2b	625 ± 69	5994 ± 210	6619 ± 213	6242
VV1t2b	348 ± 45	1392 ± 125	1741 ± 130	1680
VH0t2b	887 ± 90	8187 ± 224	9074 ± 226	9228
VH1t2b	234 ± 37	811 ± 69	1045 ± 81	1033
HH0t2b	174 ± 26	748 ± 46	921 ± 52	951
HH1t2b	10 ± 5	53 ± 11	63 ± 12	73
XX2t2b	14 ± 4	109 ± 77	123 ± 78	66
VV0t3b	96 ± 24	677 ± 63	772 ± 62	885
VV1t3b	74 ± 22	265 ± 52	339 ± 53	402
VH0t3b	183 ± 42	2037 ± 114	2220 ± 103	2185
VH1t3b	98 ± 24	483 ± 67	581 ± 68	526
HH0t3b	77 ± 28	608 ± 46	685 ± 50	724
HH1t3b	16 ± 6	83 ± 18	99 ± 19	94
XX2t3b	9 ± 4	15 ± 12	25 ± 12	45

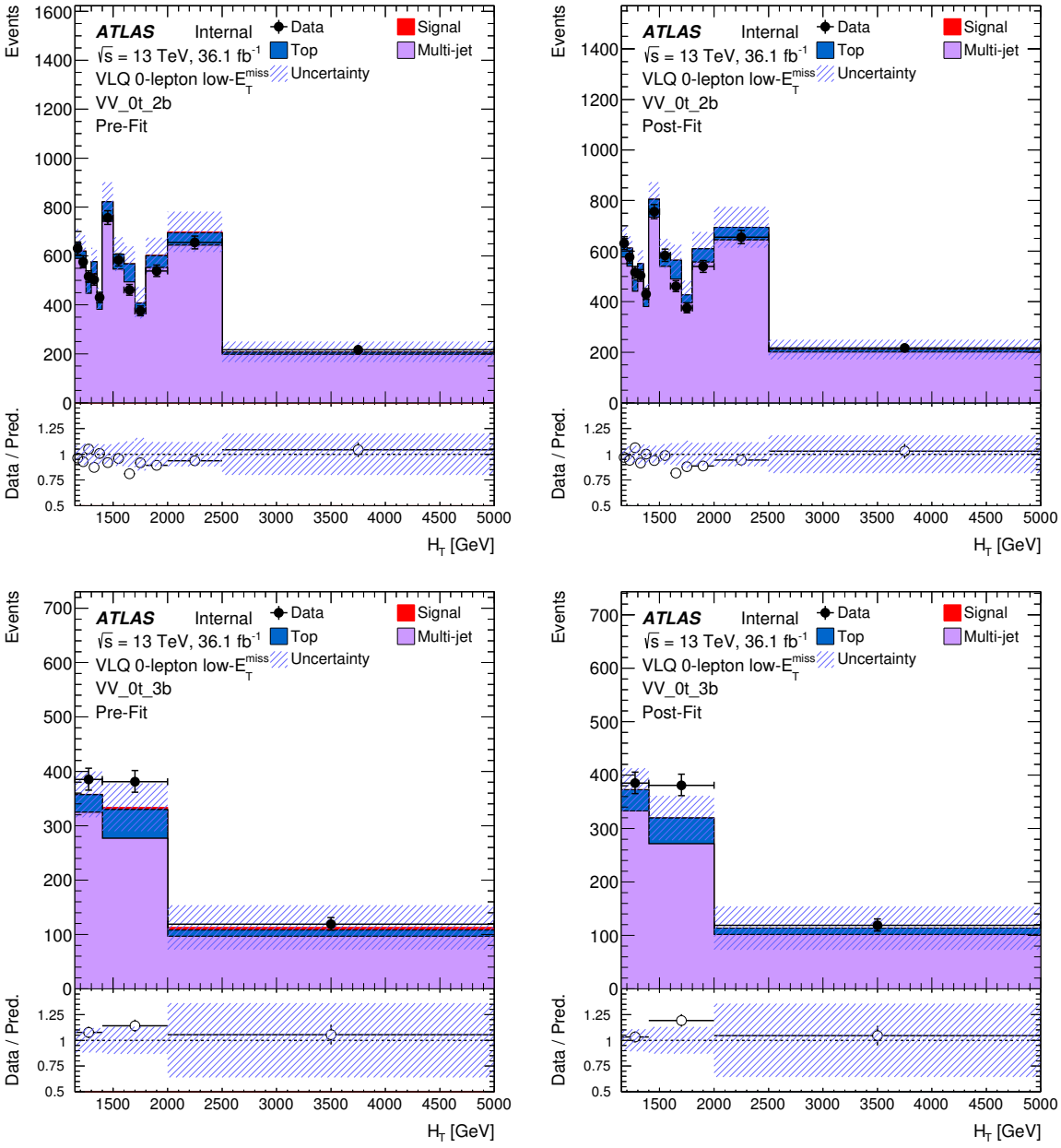

Figure 5.68: Comparison between data and prediction for the H_{T} distribution before (left) and after (right) the fit to the data under the background-only hypothesis. The figures represent the VV signal regions with 0 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) = 1) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

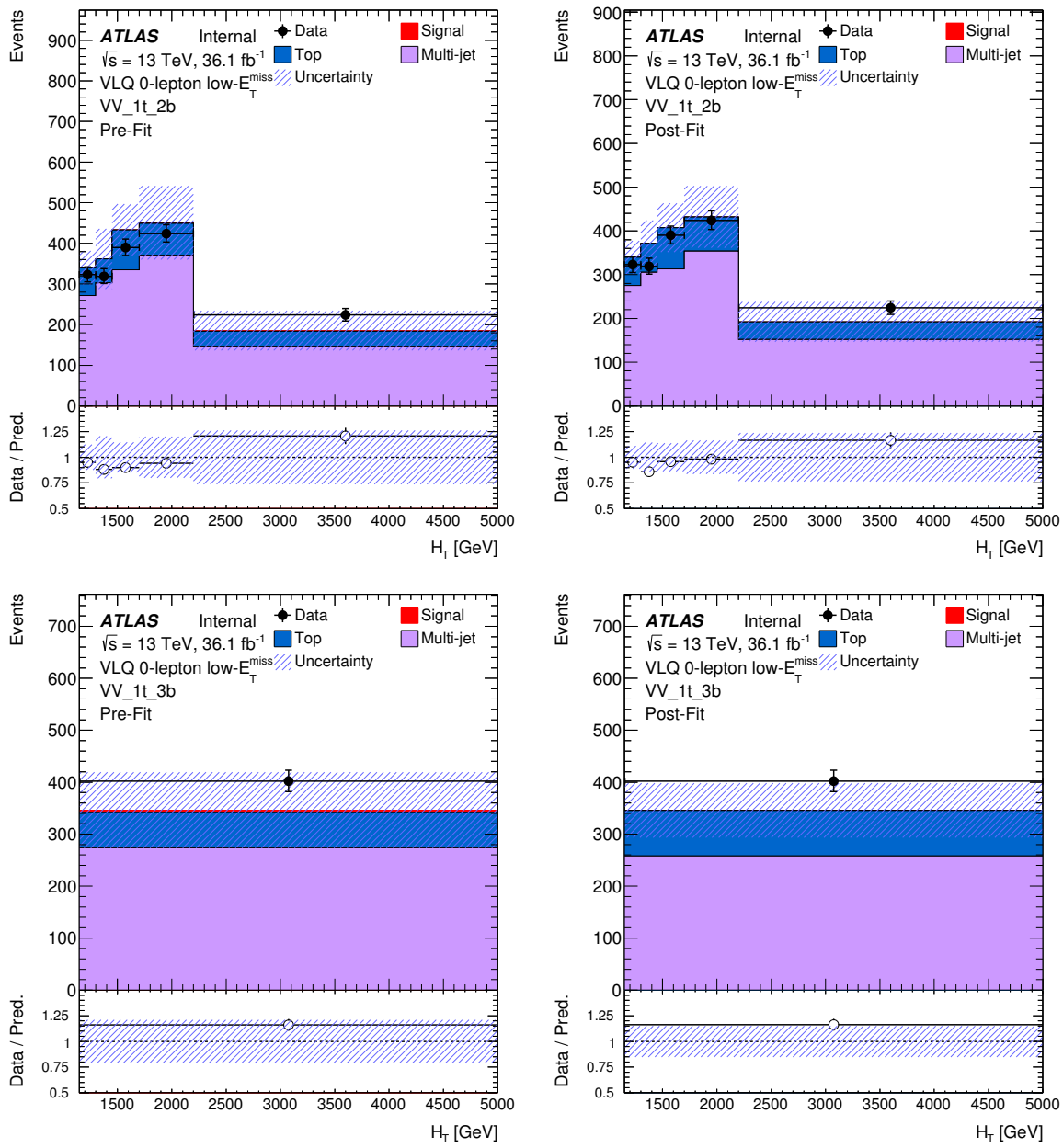

Figure 5.69: Comparison between data and prediction for the H_{T} distribution before (left) and after (right) the fit to the data under the background-only hypothesis. The figures represent the VV signal regions with 1 top tag: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) = 1) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

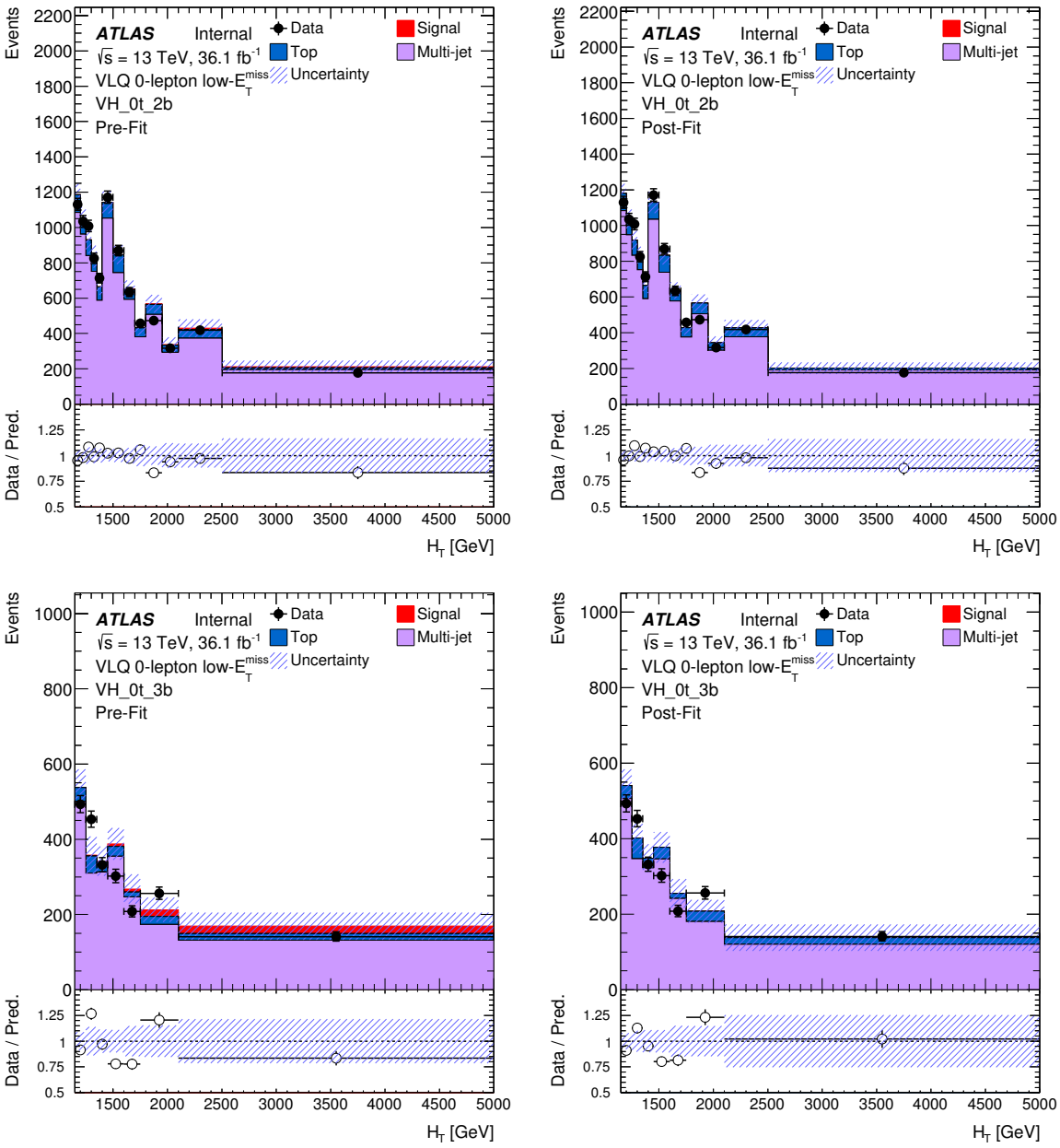

Figure 5.70: Comparison between data and prediction for the H_{T} distribution before (left) and after (right) the fit to the data under the background-only hypothesis. The figures represent the VH signal regions with 0 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) = 1) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

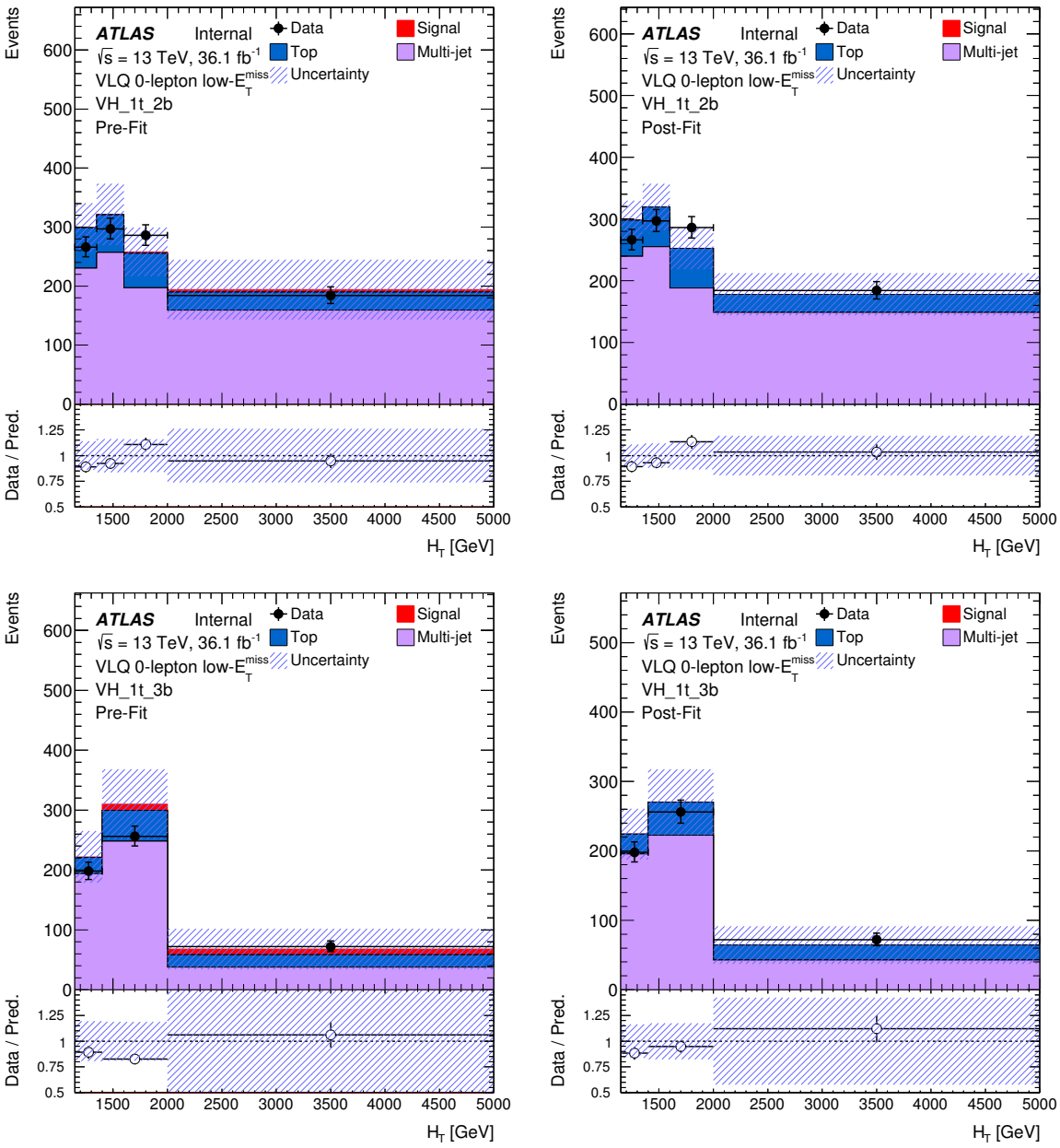

Figure 5.71: Comparison between data and prediction for the H_{T} distribution before (left) and after (right) the fit to the data under the background-only hypothesis. The figures represent the VH signal regions with 1 top tag: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

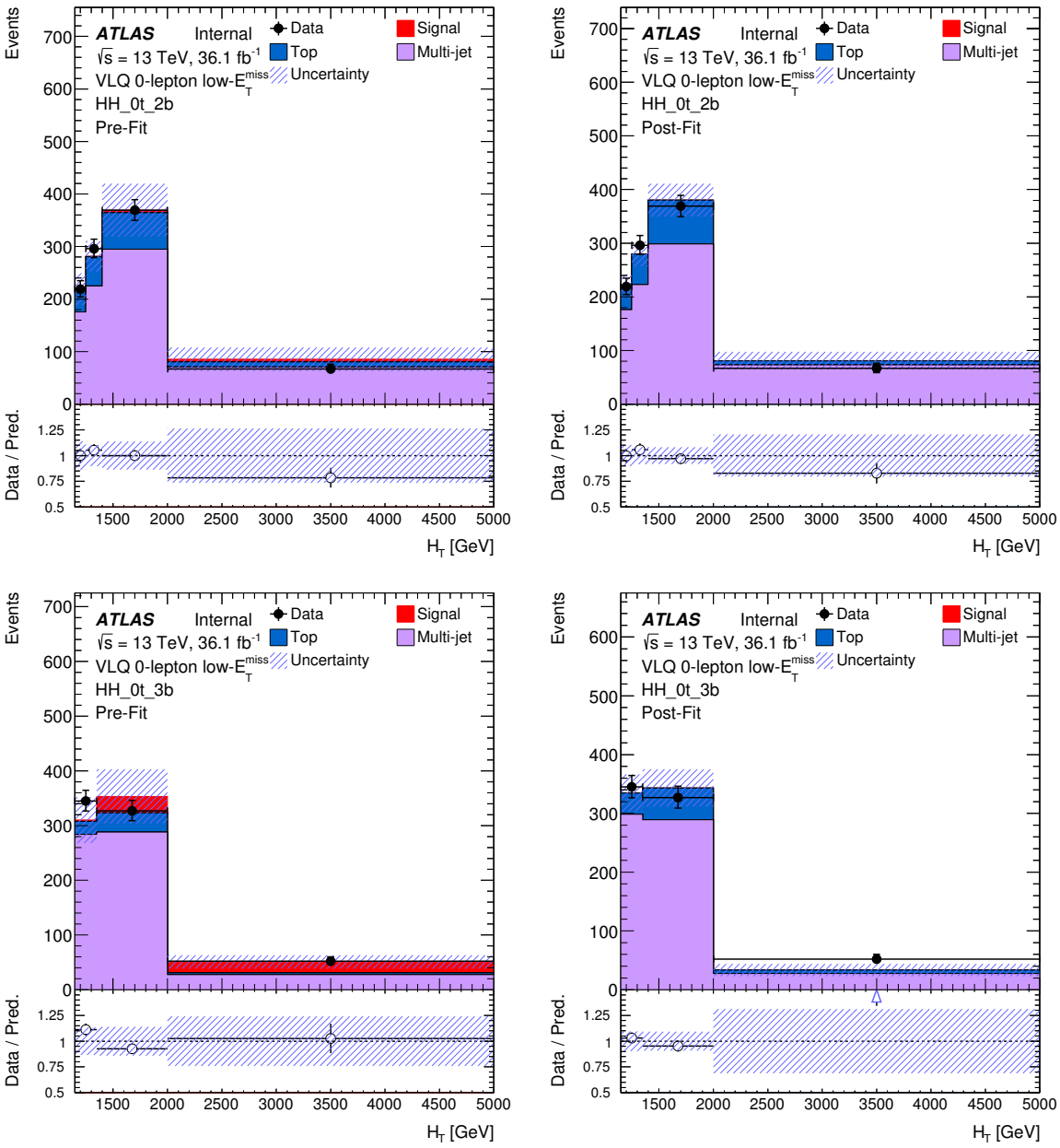

Figure 5.72: Comparison between data and prediction for the H_{T} distribution before (left) and after (right) the fit to the data under the background-only hypothesis. The figures represent the HH signal regions with 0 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) $= 1$) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

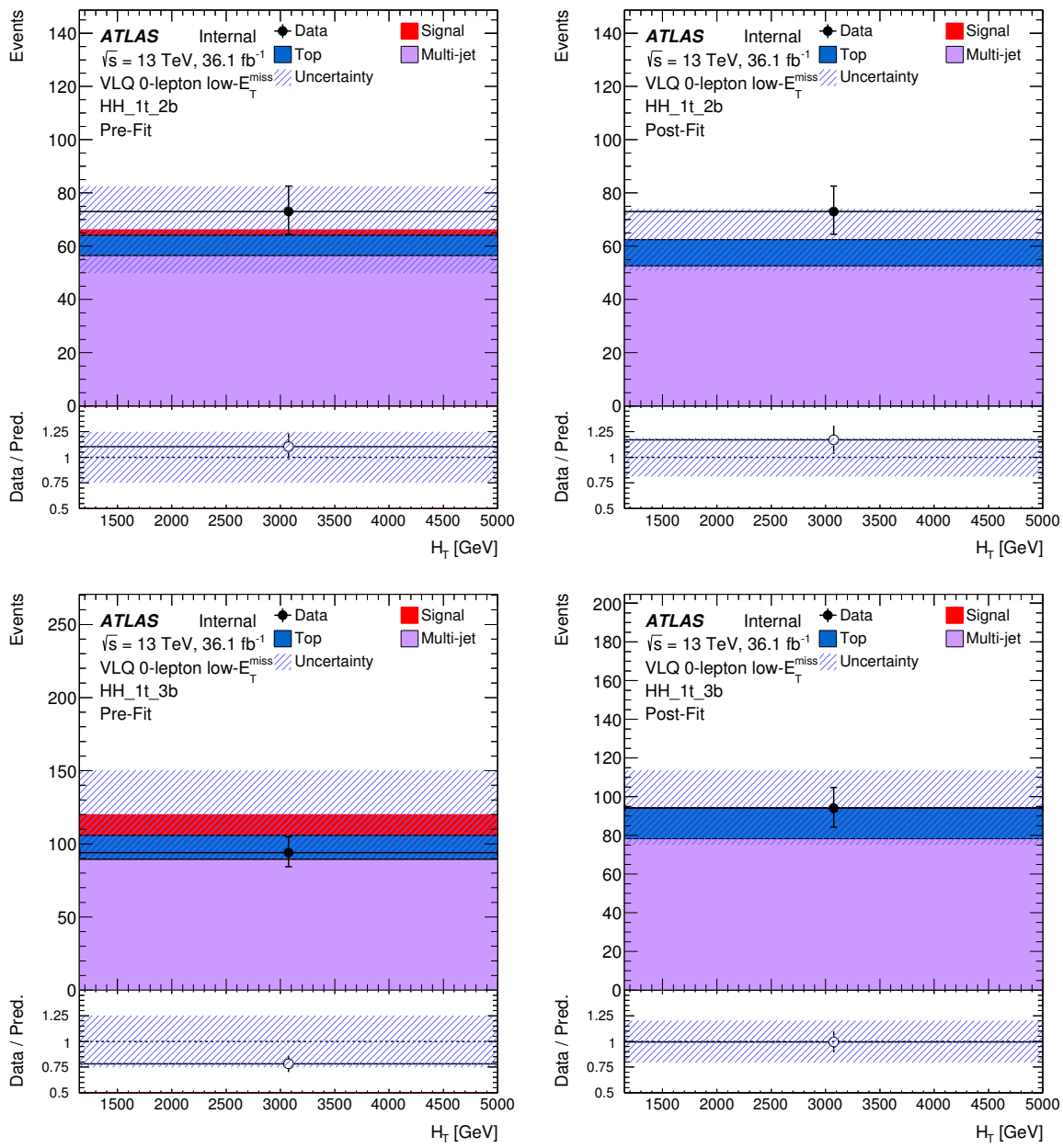

Figure 5.73: Comparison between data and prediction for the H_{T} distribution before (left) and after (right) the fit to the data under the background-only hypothesis. The figures represent the HH signal regions with 1 top tag: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) = 1) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

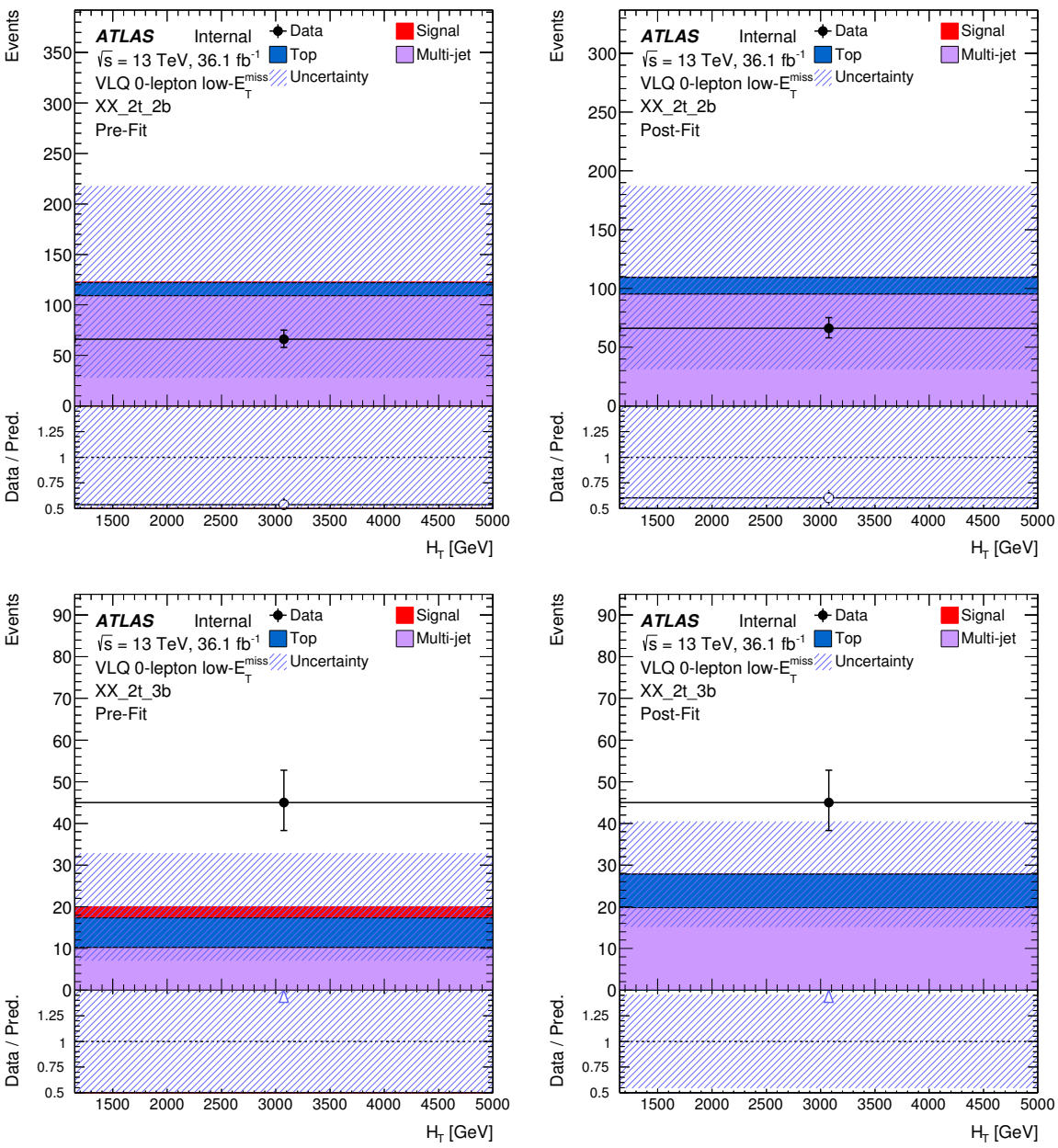

Figure 5.74: Comparison between data and prediction for the H_{T} distribution before (left) and after (right) the fit to the Asimov data under the background-only hypothesis. The figures represent the signal regions with ≥ 2 top tags: 2 b-tags (top row) and ≥ 3 b-tags (bottom row). The expected $B\bar{B}$ ($m_B = 1$ TeV; BR($B \to Hb$) = 1) signal is also shown in pre-fit plots. The hatched area represents the total uncertainty on the background.

Fig. 5.75 displays the pulls and constraints of fitted nuisance parameters. A ranking of the nuisance parameters according to the absolute post-fit effect on the fitted signal strength $\hat{\mu}$ in the fit to data for the signal-plus-background hypothesis can be seen in Fig. 5.76 for $B\bar{B}$ ($m_B = 1$ TeV; BR($T \to Hb$) = 1). The uncertainty is the only nuisance parameter pulled outside of 1σ and it is not a dominant uncertainty, i.e., it is ranked outside the top 10.

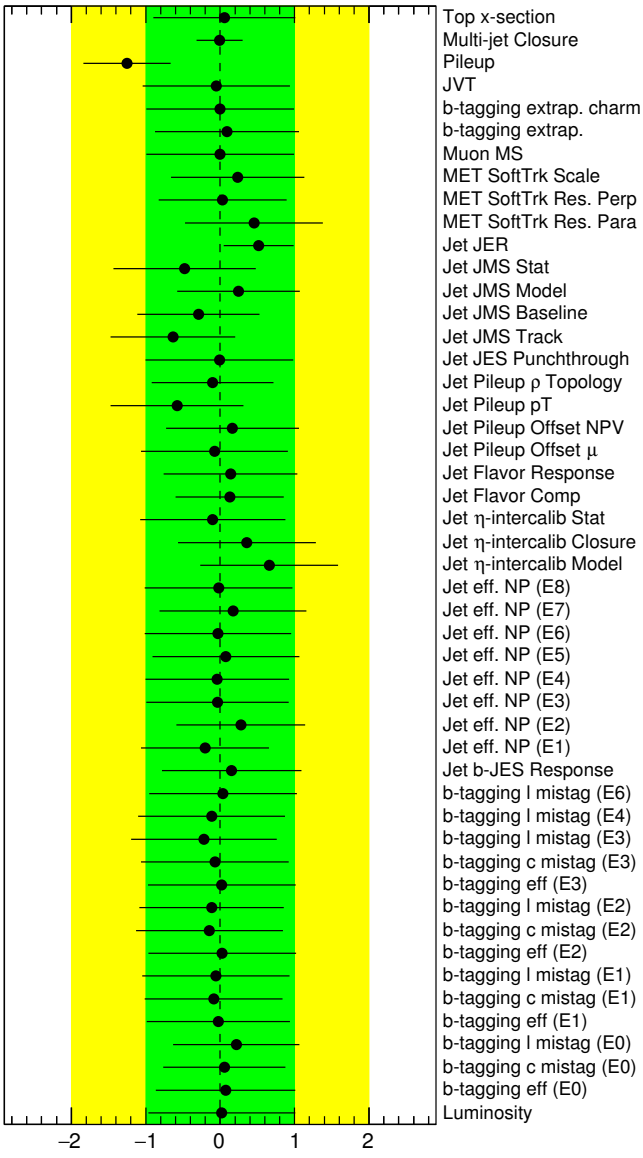

Figure 5.75: Fitted nuisance parameters in the background-only hypothesis to data. The constraints of the nuisance parameters can be read directly from the figure.

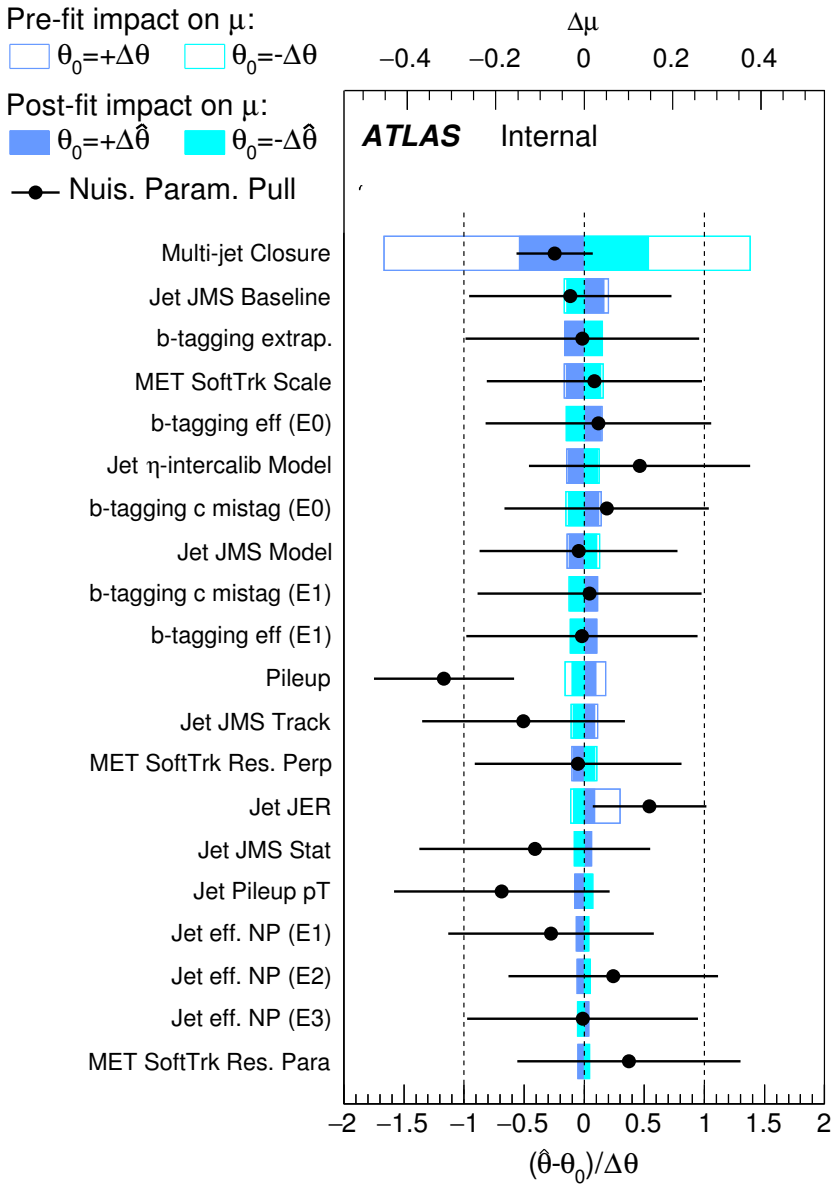

Figure 5.76: Ranking of nuisance parameters based on the fit to Asimov data in the signal-plus background hypothesis according to their effect on the uncertainty on μ ($\Delta\mu$).

Table 5.13: Expected and observed 95% CL limits on the VLQ mass for $T\bar{T}$ and $B\bar{B}$ production. Different branching ratios are presented for T and B.

Branching Ratio	Expected [GeV]	Observed [GeV]
BR($T \to Wb$) = 1	650	725
BR($T \to Zt$) = 1	750	650
BR($T \to Ht$) = 1	1001	850
BR($B \to Wt$) = 1	627	600
BR($B \to Zb$) = 1	750	713
BR($B \to Hb$) = 1	1045	903

To probe the complete branching ratio plane spanned by all decay modes, the signal samples are re-weighted and the fit is re-calculated. The complete analysis is repeated for each grid point. Fig. 5.79 shows the corresponding expected and observed 95% CL upper limits the VLQ mass for T in the BR($T \to Ht$) versus BR($T \to Wb$) plane and B in the BR($B \to Hb$) versus BR($B \to Wt$) plane, obtained by linear interpolation of the calculated CL_s versus m_T, m_B.

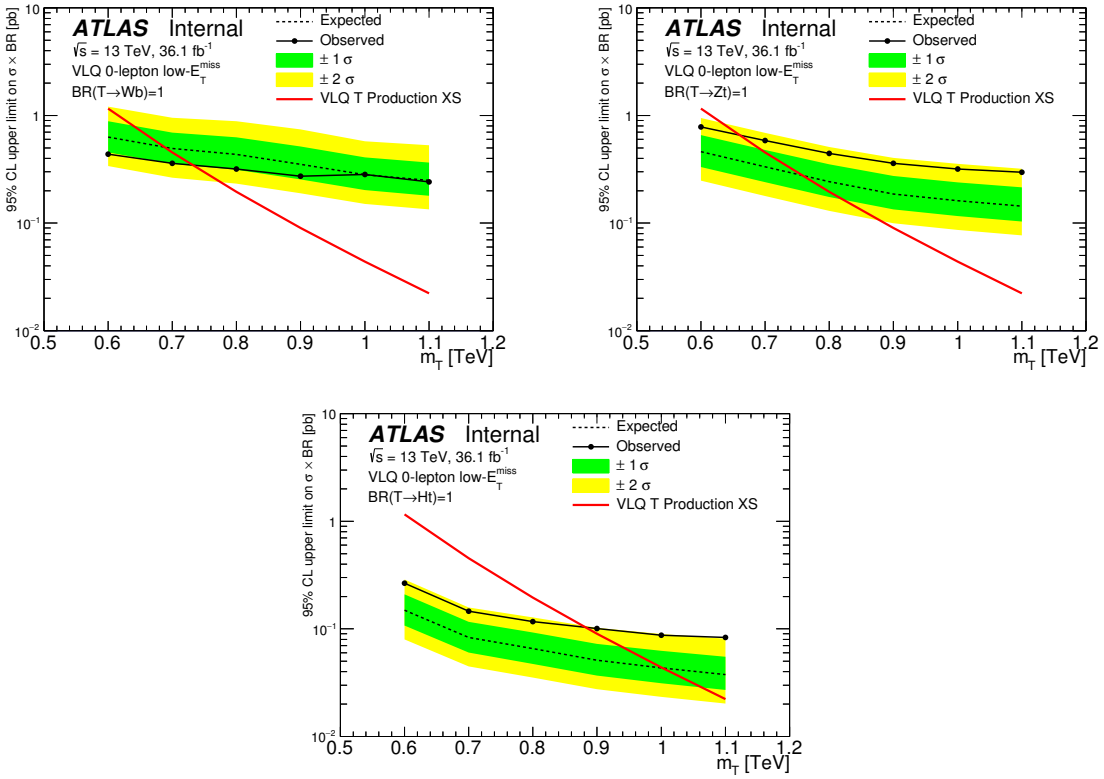

Figure 5.77: Expected (dashed black line) and observed (solid black line) upper limits at the 95% CL on the $T\bar{T}$ cross section as a function of m_T under the assumption of BR($T \to Wb$) = 1 (top-left), BR($T \to Zt$) = 1 (top-right), and BR($T \to Ht$) = 1 (bottom). The surrounding green and yellow bands correspond to ± 1 and ± 2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction.

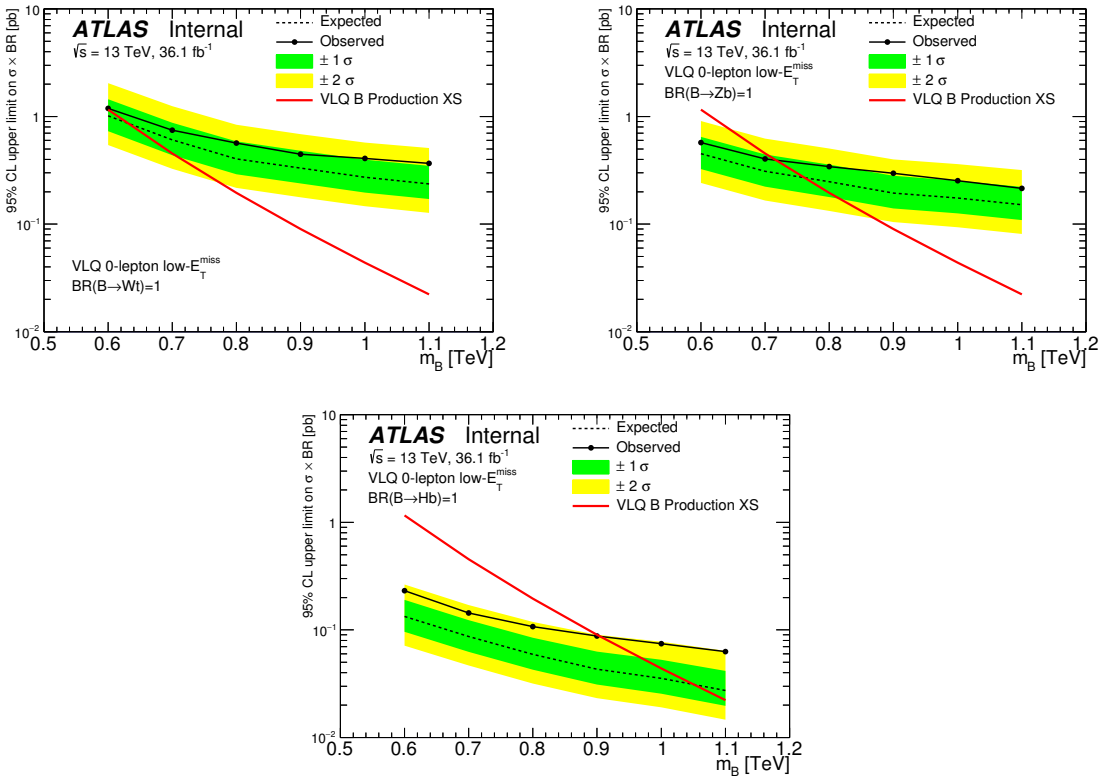

Figure 5.78: Expected (dashed black line) and observed (solid black line) upper limits at the 95% CL on the $B\bar{B}$ cross section as a function of m_B under the assumption of BR($B \to Wt$) = 1 (top-left), BR($B \to Zb$) = 1 (top-right), and BR($B \to Hb$) = 1 (bottom). The surrounding green and yellow bands correspond to ± 1 and ± 2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction.

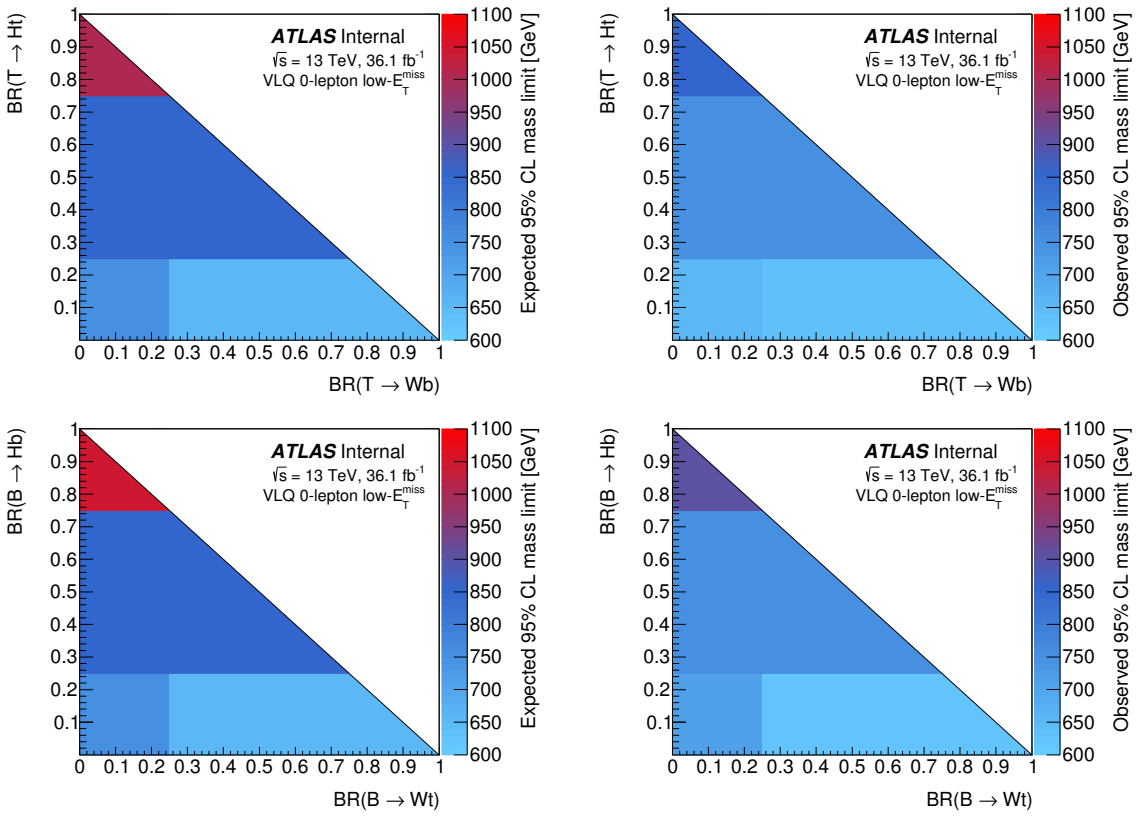

Figure 5.79: Expected (left) and observed (right) 95% CL upper limits on m_T (top) and m_B (bottom) in the branching ratio planes.

5.2.8 Conclusion

A search for pair production of vector-like quarks in the all-hadronic final state has been presented using 36.1 fb^{-1} of integrated luminosity collected by the ATLAS detector in 2015 and 2016. The analysis requires high-p_{T} small-R jets and multiple b-tags. Small-R jets are combined using a variable-R clustering algorithm and then classified with a neural network as a W/Z boson, Higgs boson, top quark, or QCD jet. The sum of all small-R jet p_{T} is used as the final discriminant across multiple categories based on the number of boson tags, top tags, and b-tags. The observed data does not have an excess consistent with the signal expectation, so a 95% CL limit is placed on VLQ pair production as a function of the hypothetical VLQ mass. Table 5.13 contains the results for the limits for VLQ pair production for the possible 'pure' branching ratios of T and B. The observed limits are lower than expected limits for each decay channel except the $\text{BR}(T \to Wb) = 1$ mode, which is the decay mode with the smallest acceptance, because of an overall excess of events in data compared to the prediction. The $T \to Ht$ and $B \to Hb$ channels are the most sensitive decay modes for $T\bar{T}$ and $B\bar{B}$ because these final states contain the most b-quarks and fall into the signal regions with a reduced contamination for multi-jet events. Additionally, limits were placed on a variety of VLQ masses across many different branching ratio combinations in Fig. 5.79. These results are competitive with recent VLQ search results in ATLAS, although slightly smaller masses are excluded for the cross sections. However, this analysis reports the strongest limit observed for the $\text{BR}(B \to Hb) = 1$ channel.

CHAPTER VI

Conclusions

VLQs are a promising signature of new physics predicted in numerous BSM theories. The VLQ mass is not tied to the electroweak scale and can be a free parameter of the model. As such, VLQs can easily be incorporated into BSM theories as they are not constrained by the recent discovery of the Higgs boson. Furthermore, VLQs belong to a larger class of 'top partners' that are expected to play a role in naturally regulating the Higgs boson mass. In theories such as the Little Higgs or Composite Higgs, VLQs emerge as a way to cancel one-loop corrections to the Higgs mass. This way, the divergence of the Higgs boson mass is postponed to a higher energy scale Λ, where Λ is specified by the BSM theory.

This dissertation presented two orthogonal searches for VLQs in 1- and 0-lepton final states with the ATLAS detector. No excess of events above the expected background were observed, and limits on the VLQ mass at the 95% CL were placed from both analyses. The performance of the 1-lepton analysis was systematically-limited where modeling uncertainties on $t\bar{t}$ production most adversely affected the sensitivity. The 0-lepton analysis was statistically-limited as the dominant background, QCD multi-jets, was estimated using a data driven technique.

While the integrated luminosities of the respective searches differ, the results can be compared to draw preliminary conclusions on the total effort to search for VLQs. (Systematically-limited analyses cannot be drastically improved by adding more data. Instead the analysis techniques can be developed further, or the systematic uncertainties must be reduced.) From the results described in Sects. 5.1.7 and 5.2.7, the two searches complement each other across the different branching ratios, considering only the $T \to Wb/Zt/Ht$ decays and compared directly in Fig. 6.1. The two figures also differ because a smaller branching ratio step-size was considered in the 1-lepton results than the 0-lepton results presented here.

The two analyses are most sensitive in opposite corners of the branching ratio plane. This improves the overall sensitivity to VLQs in the ATLAS program, near $m_T = 1$ TeV in each corner, because the full plane can be covered using just these two analyses. The

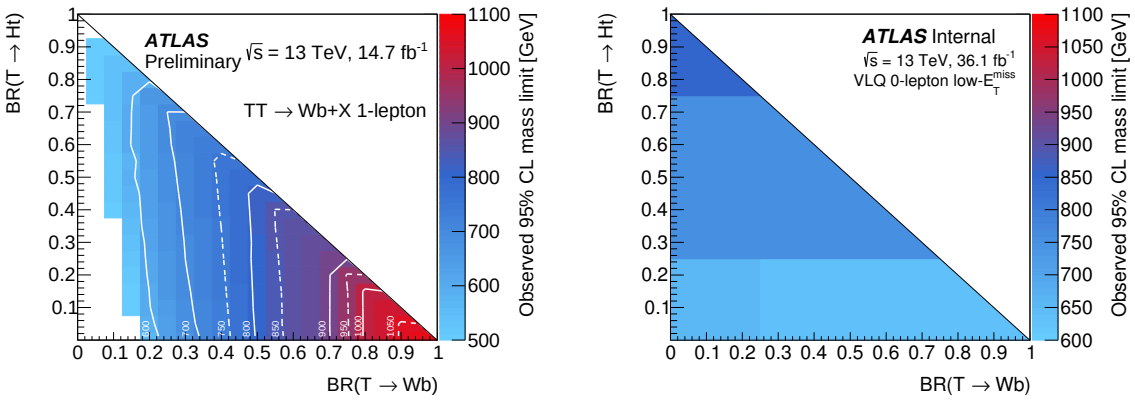

Figure 6.1: Comparison between the 95% CL limits on m_T for the 1-lepton VLQ search (left) and 0-lepton VLQ search (right).

corner $T \to Zt$ is noticeably less sensitive to the other corners due to the large multi-jet contribution in the 0-lepton analysis and the 1-lepton analysis assumes the lepton is produced by a W boson from the VLQ decay. VLQs can still be compatible with many BSM theories for masses $\lesssim$ 2 TeV and the searches at ATLAS and CMS should continue until this approximate range is excluded. However, the limits for m_T and m_B have reached the TeV scale and it is beneficial to begin considering VLQs that have more exotic decay modes than described in this dissertation [117]. In those scenarios, these results are not directly applicable because a non-negligible fraction of the VLQ decays are not considered. Furthermore, the limits on the VLQ mass are becoming large enough that pair production is no longer a favored mechanism. Instead, the preferred production mode will be single VLQs for future analyses. With more data delivered by the LHC and collected by the ATLAS experiment, more sophisticated analyses can search for VLQs in unique final states and assuming single production and both standard and more exotic decay modes. In the event of a discovery in a VLQ search, testing the signal hypothesis, to measure the signal strength, and quantifying the branching ratios will be the most critical steps to confirming the VLQ existence and directing future searches towards specific BSM theories.

APPENDICES

APPENDIX A

Multi-Classification of Variable-R Re-clustered Jets with a Deep Neural Network

Neural networks (NN) are common in b-tagging, track reconstruction, and other aspects of experimental particle physics, see Sect. 4.2. New efforts are underway within the collider physics community to exploit the power of machine learning techniques, e.g., boosted decision trees and NNs, and early analyses already show promise in discrimination power over existing techniques, particularly in the realm of object identification, for examples see [118–121]. Overviews of NNs, terminology, and development can be found in [122–124].

In this analysis, a multi-classification NN is used to distinguish vRC jets that originated from W/Z ($\equiv V$), Higgs, top, or light-jets. Similar results have been described previously [125], and this work differentiates from the existing literature by applying the NN identification to vRC jets, a novel reconstructed physics object. The NN is optimized using VLQ samples as sources of signal jets (W, Z, Higgs, and top) and multi-jet MC as a source of background jets ('light'). Prior to training the network, the multi-jet MC p_{T} distribution is re-weighted to match the signal (VLQ decays) as in Fig. A.1. Once the input files have been processed, the Keras [126] software package is used to build and train the NN with the Theano backend [127].

Fully-contained signal jets, those that have all quarks from the heavy object decay geometrically-matched to the constituents of the vRC jets[1], are used in the training with the following seventeen inputs:

- vRC jet mass

[1]To be matched, the partons i are required to be within $\Delta R(i,j) < 0.3$ of at least one subjet j of the vRC jet.

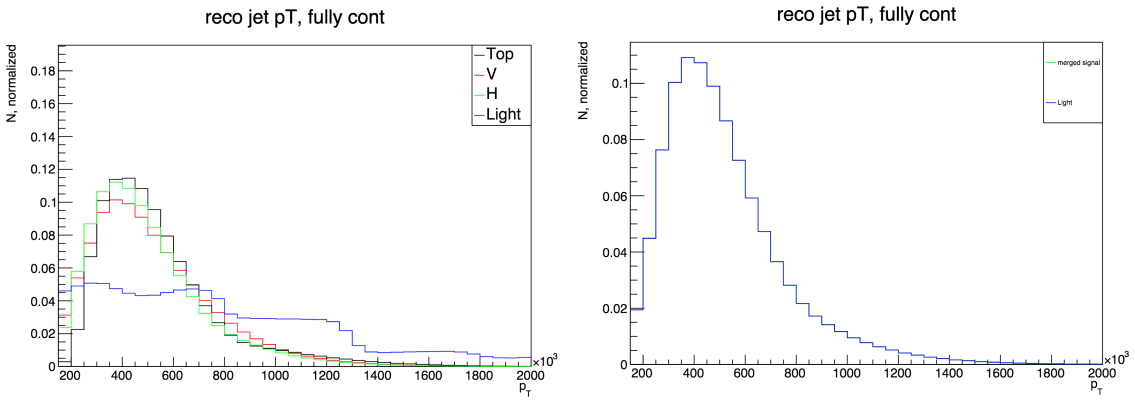

Figure A.1: Distribution of vRC jet p_{T} for signal and background before (left) and after (right) re-weighting the background p_{T} distribution using a one-dimensional scale factor.

- vRC jet number of constituents
- Kinematic information of three leading p_{T} subjets:
  - p_{T}
  - E
  - η
  - ϕ
  - b-tagging bin

The b-tagging bin is used as an input rather than the value of `mv2c10` because only specific working points, see Sect. 4.2, are calibrated for analysis use. The bin value used in the training matches the thresholds for each of the 4 working points, 0.0 (not b-tagged), 0.1758475 (60%), 0.645925 (70%), 0.8244273 (77%), and 0.934906 (85%). For vRC jets that do not have three constituents, default values are assigned:

- $p_{\mathrm{T}} = 0$
- $E = 0$
- $\eta =$ vRC jet η
- $\phi =$ vRC jet ϕ
- b-tagging bin = -1

The NN has the following architecture (similar to that applied in [118]), with terms explicitly defined in Keras:

- 7 Fully connected layers ("Dense")
- Number of nodes (5 hidden layers): 30, 25, 23, 13, 11
  - 17 inputs
  - 4 outputs: light, W/Z, H, top
  - Activation function: rectified linear units (hidden layers) and sigmoid (output layer)
- Optimizer: Adam [128]
- Learning rate: 0.0001
- L1 Regularizer: 0.01
- Batch size: 100 (Batch Normalization)
- Trained for 100 epochs with early stopping

Once the network is trained, the network is tested on an orthogonal dataset to ensure there was no overtraining. The NN output is converted to a probability for each jet to maximize the light jet rejection, detailed in Eq. 5.16 and shown in Fig. A.2. The performance of the NN tagging as a function of vRC jet p_{T} is shown in Fig. A.3.

A selection on the probability result defines a tagging efficiency. This analysis assumes a selection that performs at 70% efficient for V, 70% efficient for Higgs, and 50% efficient for top. After the baseline selection and requiring at least two b-tags with $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV, the measured efficiencies in simulated events (not requiring full containment of truth partons) are 66% (V), 52% (Higgs), and 46% (top).

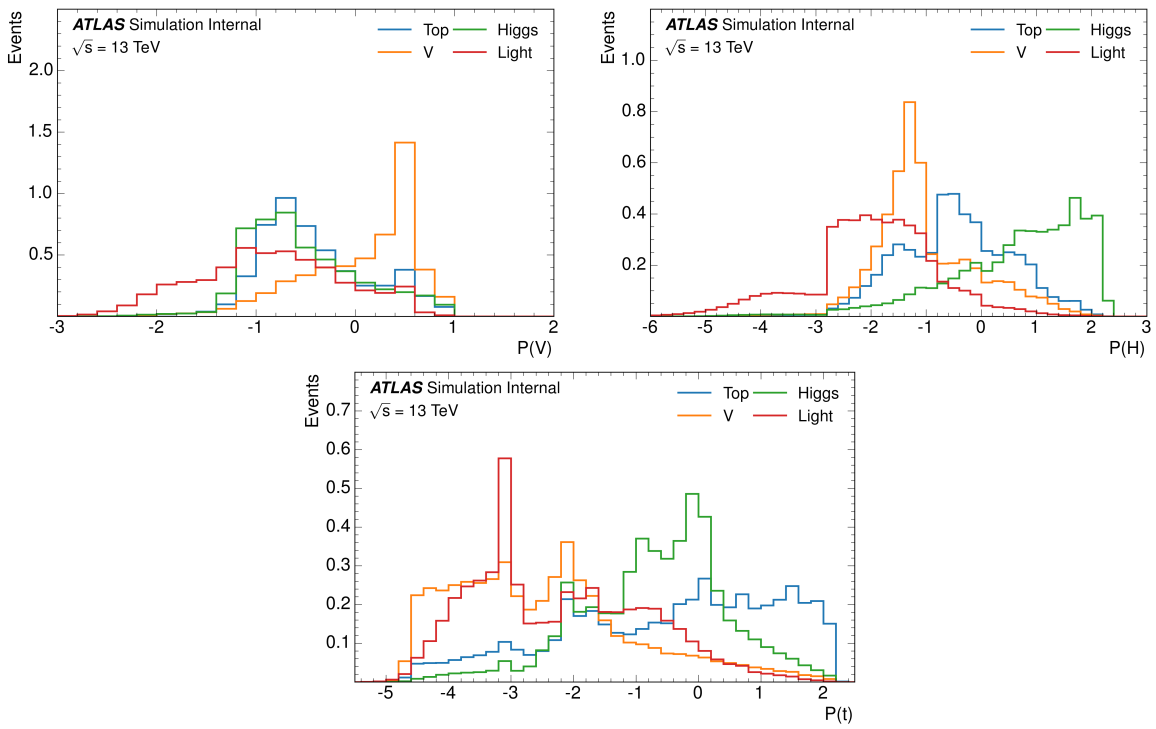

Figure A.2: Probability distributions for vRC jets NN output: $P(V)$ (top-left), $P(H)$ (top-right), and $P(t)$ (bottom). Each distribution shows the probability for jets from all sources (V, H, top, and light).

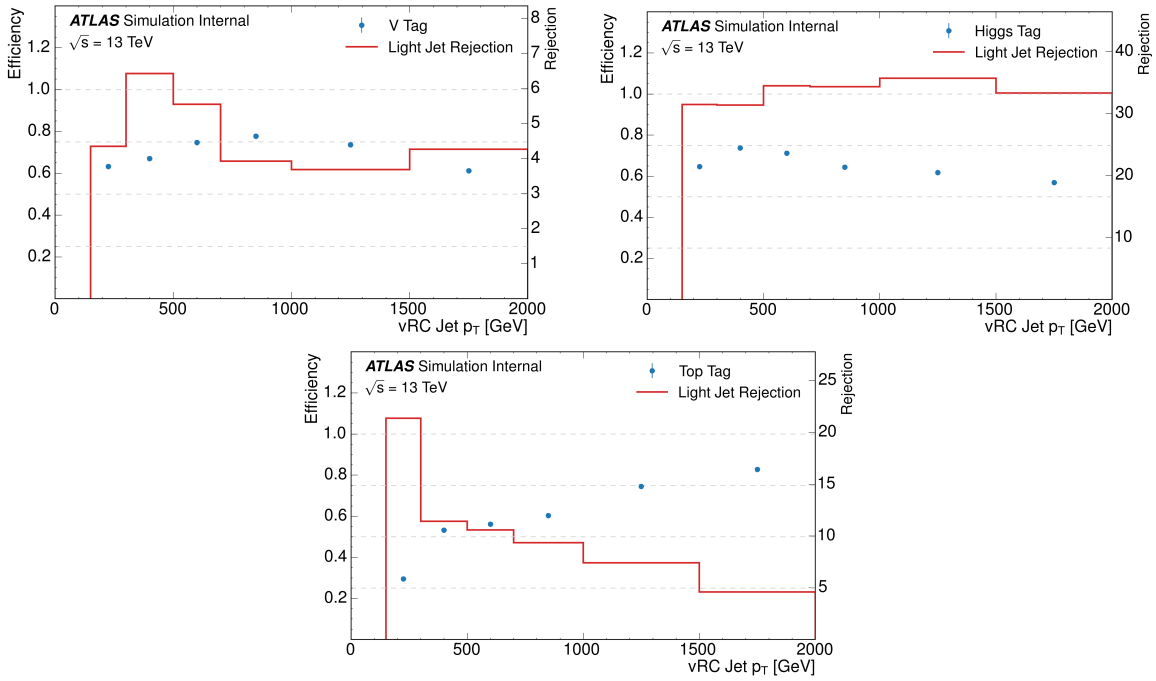

Figure A.3: Tagging efficiency and light jet rejection for vRC jets identified as truth $V = W/Z$ (top-left), Higgs (top-right), top (bottom).

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[2] ATLAS Collaboration Collaboration,, *Search for new phenomena in $t\bar{t}$ final states with additional heavy-flavour jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-104, CERN, Geneva, Sep, 2016. https://cds.cern.ch/record/2220371.

[3] ATLAS Collaboration Collaboration,, *Search for pair production of vector-like top quarks in events with one lepton and an invisibly decaying Z boson in $\sqrt{s} = 13$ TeV pp collisions at the ATLAS detector*, Tech. Rep. ATLAS-CONF-2017-015, CERN, Geneva, Mar, 2017. https://cds.cern.ch/record/2257730.

[4] ATLAS Collaboration Collaboration,, *Search for single production of vector-like quarks decaying into Wb in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-072, CERN, Geneva, Aug, 2016. https://cds.cern.ch/record/2206226.

[5] ATLAS Collaboration Collaboration,, *Search for new physics using events with b-jets and a pair of same charge leptons in 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-032, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2161545.

[6] CMS Collaboration Collaboration,, *Search for single production of vector-like quarks decaying to a Z boson and a top or a bottom quark in proton-proton collisions at 13 TeV*, Tech. Rep. CMS-PAS-B2G-17-007, CERN, Geneva, 2017. https://cds.cern.ch/record/2256762.

[7] CMS Collaboration Collaboration,, *Search for heavy vector-like quarks decaying to same-sign dileptons*, Tech. Rep. CMS-PAS-B2G-16-019, CERN, Geneva, 2017. https://cds.cern.ch/record/2256747.

[8] CMS Collaboration Collaboration,, *Search for exotic light-flavor quark partners in pp collisions at $\sqrt{s} = 8$* TeV, Tech. Rep. CMS-PAS-B2G-12-016, CERN, Geneva, 2016. https://cds.cern.ch/record/2235750.

[9] CMS Collaboration Collaboration,, *Search for vector-like quark pair production in final states with leptons and boosted Higgs bosons at $\sqrt{s} = 13$* TeV, Tech. Rep.

CMS-PAS-B2G-16-011, CERN, Geneva, 2016.
https://cds.cern.ch/record/2220829.

[10] CMS Collaboration Collaboration,, *Search for a heavy resonance decaying to a top quark and a vector-like top quark at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-B2G-16-013, CERN, Geneva, Mar, 2017. https://cds.cern.ch/record/2256071. Submitted to JHEP. All figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/B2G-16-013.

[11] CMS Collaboration Collaboration,, *Search for single production of vector-like quarks decaying into a b quark and a W boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-B2G-16-006. CMS-B2G-16-006, CERN, Geneva, Jan, 2017. https://cds.cern.ch/record/2243273. Submitted to Phys. Lett. B. All figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/B2G-16-006.

[12] CMS Collaboration Collaboration,, *Search for single production of vector-like quarks decaying to a Z boson and a top or a bottom quark in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Tech. Rep. CERN-EP-2016-326. CMS-B2G-16-001, CERN, Geneva, Jan, 2017. https://cds.cern.ch/record/2243017. Submitted to JHEP. All figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/B2G-16-001.

[13] CMS Collaboration Collaboration,, *Search for electroweak production of a vector-like quark decaying to a top quark and a Higgs boson using boosted topologies in fully hadronic final states*, Tech. Rep. CMS-B2G-16-005. CMS-B2G-16-005, CERN, Geneva, Dec, 2016. https://cds.cern.ch/record/2239298. Submitted to JHEP. All figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/B2G-16-005.

[14] CMS Collaboration Collaboration,, *Search for single production of a heavy vector-like T quark decaying to a Higgs boson and a top quark with a lepton and jets in the final state*, Tech. Rep. CERN-EP-2016-279. CMS-B2G-15-008, CERN, Geneva, Dec, 2016. https://cds.cern.ch/record/2237477. Submitted to Phys. Lett. B.

[15] CMS Collaboration Collaboration,, *Search for pair production of vector-like T quarks in the lepton plus jets final state*, Tech. Rep. CMS-PAS-B2G-16-002, CERN, Geneva, 2016. https://cds.cern.ch/record/2141070.

[16] CMS Collaboration Collaboration,, *Search for top quark partners with charge 5/3 at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-PAS-B2G-15-006, CERN, Geneva, 2015. https://cds.cern.ch/record/2114805.

[17] ATLAS Collaboration Collaboration, G. Aad et al., Phys.Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[18] CMS Collaboration, S. Chatrchyan et al., Phys. Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[19] G. Kane, *Modern Elementary Particle Physics: The Fundamental Particles and Forces?* Advanced book classics. Avalon Publishing, 1993. https://books.google.com/books?id=3oC-2BiV7AsC.

[20] M. Srednicki, *Quantum Field Theory*. Cambridge University Press, 2007. https://books.google.com/books?id=5OepxIG42B4C.

[21] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version*. Physics textbook. Wiley, New York, NY, 2008. https://cds.cern.ch/record/111880.

[22] T. Lancaster, S. Blundell, and S. Blundell, *Quantum Field Theory for the Gifted Amateur*. OUP Oxford, 2014. https://books.google.com/books?id=Y-0kAwAAQBAJ.

[23] D. Galbraith, *UX: Standard Model of the Standard Model*, http://davidgalbraith.org/portfolio/ux-standard-model-of-the-standard-model/. Accessed: 02-2017.

[24] A. Djouadi and A. Lenz, Phys. Lett. **B715** (2012) 310–314, arXiv:1204.1252 [hep-ph].

[25] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, U. Nierste, and M. Wiebusch, Phys. Rev. Lett. **109** (2012) 241802, arXiv:1209.1101 [hep-ph].

[26] Particle Data Group Collaboration, C. Patrignani et al., Chin. Phys. **C40** no. 10, (2016) 100001.

[27] UA1 Collaboration Collaboration,, Physics Letters B **122** no. 1, (1983) 103 – 116.

[28] UA1 Collaboration Collaboration,, Physics Letters B **126** no. 5, (1983) 398 – 410.

[29] UA2 Collaboration Collaboration,, Physics Letters B **122** no. 5, (1983) 476 – 485.

[30] UA2 Collaboration Collaboration,, Physics Letters B **129** no. 1, (1983) 130 – 140.

[31] P. Higgs, Physics Letters **12** no. 2, (1964) 132 – 133, http://www.sciencedirect.com/science/article/pii/0031916364911369.

[32] F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964) 321–323, https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[33] P. W. Higgs, Phys. Rev. Lett. **13** (1964) 508–509, https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[34] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Lett. **13** (1964) 585–587, https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[35] P. W. Higgs, Phys. Rev. **145** (1966) 1156–1163, https://link.aps.org/doi/10.1103/PhysRev.145.1156.

[36] T. W. B. Kibble, Phys. Rev. **155** (1967) 1554–1561, https://link.aps.org/doi/10.1103/PhysRev.155.1554.

[37] SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3 Collaboration, S. Schael et al., Phys. Rept. **427** (2006) 257–454, arXiv:hep-ex/0509008 [hep-ex].

[38] N. Arkani-Hamed, A. G. Cohen, E. Katz, and A. E. Nelson, JHEP **07** (2002) 034, arXiv:hep-ph/0206021 [hep-ph].

[39] M. Schmaltz and D. Tucker-Smith, Ann. Rev. Nucl. Part. Sci. **55** (2005) 229–270, arXiv:hep-ph/0502182 [hep-ph].

[40] D. B. Kaplan, H. Georgi, and S. Dimopoulos, Phys. Lett. **B136** (1984) 187–190.

[41] K. Agashe, R. Contino, and A. Pomarol, Nucl. Phys. **B719** (2005) 165–187, arXiv:hep-ph/0412089 [hep-ph].

[42] F. del Aguila, M. Pérez-Victoria, and J. Santiago, JHEP **09** (2000) 011, arXiv:hep-ph/0007316 [hep-ph].

[43] J. A. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Pérez-Victoria, Phys. Rev. **D88** no. 9, (2013) 094010, arXiv:1306.0572 [hep-ph].

[44] M. Czakon and A. Mitov, Comput. Phys. Commun. **185** (2014) 2930, arXiv:1112.5675 [hep-ph].

[45] F. del Aguila, L. Ametller, G. Kane, and J. Vidal, Nuclear Physics B **334** no. 1, (1990) 1–23.

[46] J. A. Aguilar-Saavedra, JHEP **11** (2009) 030, arXiv:0907.3155 [hep-ph].

[47] J. Aguilar-Saavedra, *Protos: PROgram for TOp Simulations*, tech. rep. http://jaguilar.web.cern.ch/jaguilar/protos/.

[48] ATLAS Collaboration Collaboration,, *Search for pair production of heavy vector-like quarks decaying to high-p_T W bosons and b quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-102, CERN, Geneva, Sep, 2016. https://cds.cern.ch/record/2219436.

[49] L. Evans and P. Bryant, Journal of Instrumentation **3** no. 08, (2008) S08001, http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[50] T. A. Collaboration, Journal of Instrumentation **3** no. 08, (2008) S08003, http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[51] *ATLAS Coordinate System*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/CcordinateSystemFootnote. Accessed: 2017-01-21.

[52] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, Sep, 2010. https://cds.cern.ch/record/1291633.

[53] ATLAS Collaboration, M. Aaboud et al., arXiv:1611.09661 [hep-ex].

[54] M. L. Mangano and T. J. Stelzer, Ann. Rev. Nucl. Part. Sci. **55** (2005) 555–588.

[55] ATLAS Collaboration, G. Aad et al., Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[56] GEANT4 Collaboration, S. Agostinelli et al., Nucl. Instrum. Meth. **A506** (2003) 250–303.

[57] *Early Inner Detector Tracking Performance in the 2015 data at $\sqrt{s}$ = 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-051, CERN, Geneva, Dec, 2015. https://cds.cern.ch/record/2110140.

[58] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger, Journal of Physics: Conference Series **119** no. 3, (2008) 032014, http://stacks.iop.org/1742-6596/119/i=3/a=032014.

[59] ATLAS Collaboration Collaboration,, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2157687.

[60] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008.

[61] ATLAS Collaboration, G. Aad et al., Eur. Phys. J. **C76** no. 5, (2016) 292, arXiv:1603.05598 [hep-ex].

[62] ATLAS Collaboration Collaboration,, *Photon identification in 2015 ATLAS data*, Tech. Rep. ATL-PHYS-PUB-2016-014, CERN, Geneva, Aug, 2016. https://cds.cern.ch/record/2203125.

[63] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, Tech. Rep. ATL-PHYS-PUB-2015-045, CERN, Geneva, Nov, 2015. https://cds.cern.ch/record/2064383.

[64] ATLAS Collaboration, G. Aad et al., arXiv:1603.02934 [hep-ex].

[65] P. Speckmayer, T. Carli, and C. W. Fabjan, *Energy Measurement of Hadrons with the CERN ATLAS Calorimeter*. PhD thesis, Vienna, Tech. U., Vienna, 2008. https://cds.cern.ch/record/1112036. Presented on 18 Jun 2008.

[66] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[67] M. Cacciari and G. P. Salam, Phys. Lett. **B641** (2006) 57–61, arXiv:hep-ph/0512210 [hep-ph].

[68] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[69] ATLAS Collaboration, M. Aaboud et al., arXiv:1703.09665 [hep-ex].

[70] M. Cacciari and G. P. Salam, Phys. Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[71] ATLAS Collaboration, G. Aad et al., Eur. Phys. J. **C76** no. 11, (2016) 581, arXiv:1510.03823 [hep-ex].

[72] A. Collaboration, JHEP **1309** (2013) 076, arXiv:1306.4945 [hep-ex].

[73] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. Webber, Phys.Lett. **B269** (1991) 432–438.

[74] S. D. Ellis and D. E. Soper, Phys.Rev. **D48** (1993) 3160–3166, arXiv:hep-ph/9305266 [hep-ph].

[75] A. Collaboration, *Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS*, Tech. Rep. ATL-PHYS-PUB-2015-033, 2015. https://cds.cern.ch/record/2041461.

[76] ATLAS Collaboration Collaboration,, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2160731.

[77] *Expected performance of the ATLAS b-tagging algorithms in Run-2*, Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037697.

[78] *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2037904.

[79] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[80] *ATLAS Run 1 Pythia8 tunes*, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014. https://cds.cern.ch/record/1966419.

[81] M. Botje et al., arXiv:1101.0538 [hep-ph].

[82] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[83] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, P. Nadolsky, J. Pumplin, D. Stump, and C. P. Yuan, Phys. Rev. **D89** no. 3, (2014) 033009, arXiv:1302.6246 [hep-ph].

[84] R. D. Ball et al., Nucl. Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[85] S. Alioli, P. Nason, C. Oleari, and E. Re, JHEP **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[86] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna, and E. Norrbin, Comput. Phys. Commun. **135** (2001) 238–259, arXiv:hep-ph/0010017 [hep-ph].

[87] P. Z. Skands, Phys. Rev. **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[88] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, JHEP **01** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[89] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[90] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[91] GEANT4 Collaboration, S. Agostinelli et al., Nucl. Instrum. Meth. **A506** (2003) 250–303.

[92] A. Collaboration, *Boosted hadronic top identification at ATLAS for early 13 TeV data*, Tech. Rep. ATL-PHYS-PUB-2015-053, 2015. https://cds.cern.ch/record/2116351.

[93] A. J. Larkoski, I. Moult, and D. Neill, JHEP **12** (2014) 009, arXiv:1409.6298 [hep-ph].

[94] A. J. Larkoski, I. Moult, and D. Neill, JHEP **05** (2016) 117, arXiv:1507.03018 [hep-ph].

[95] J. Thaler and K. Van Tilburg, JHEP **03** (2011) 015, arXiv:1011.2268 [hep-ph].

[96] J. Thaler and K. Van Tilburg, JHEP **02** (2012) 093, arXiv:1108.2701 [hep-ph].

[97] ATLAS Collaboration,, JHEP **08** (2015) 105, arXiv:1505.04306 [hep-ex].

[98] J. Butterworth, E. Dobson, U. Klein, B. Mellado Garcia, T. Nunnemann, J. Qian, D. Rebuzzi, and R. Tanaka, *Single Boson and Diboson Production Cross Sections in pp Collisions at sqrts=7 TeV*, Tech. Rep. ATL-COM-PHYS-2010-695, CERN, Geneva, Aug, 2010. https://cds.cern.ch/record/1287902.

[99] Atlas Collaboration Collaboration, G. Aad et al., Eur.Phys.J. **C71** (2011) 1577, arXiv:1012.1792 [hep-ex].

[100] K. Becker, T. Cornelissen, F. Derue, A. Henrichs, D. Hirschb, X. Lei, O. Nackenhorst, F. O'Grady, D. Pelikan, M. Pinamonti, S. Pires, J. Sjin, and P. Tepel, *Estimation of Fake Lepton Background for Top Analyses Using the* $\sqrt{s} = 8$ *TeV Dataset*, Tech. Rep. ATL-COM-PHYS-2013-1100, CERN, Geneva, Aug, 2013.

[101] ATLAS Collaboration, M. Aaboud et al., Eur. Phys. J. **C76** no. 12, (2016) 653, arXiv:1608.03953 [hep-ex].

[102] K. Melnikov and F. Petriello, Phys. Rev. **D74** (2006) 114017, arXiv:hep-ph/0609070 [hep-ph].

[103] J. M. Campbell and R. K. Ellis, Phys. Rev. **D60** (1999) 113006, arXiv:hep-ph/9905386 [hep-ph].

[104] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer, Comput. Phys. Commun. **191** (2015) 74–89, arXiv:1406.4403 [hep-ph].

[105] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, Comput. Phys. Commun. **182** (2011) 1034–1046, arXiv:1007.1327 [hep-ph].

[106] W. Verkerke and D. Kirkby, arXiv:0306116 [physics].

[107] W. Verkerke and D. Kirkby, http://rootfit.sourceforge.net.

[108] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J C **71** (2011) 1554, arXiv:1007.1727 [physics.data-an].

[109] A. L. Read, J. Phys. G **28** (2002) 2693.

[110] T. Junk, Nucl. Instrum. Meth. **434** (1999) 435, arXiv:9902006 [hep-ex].

[111] J. H. Friedman, *Data analysis techniques for high energy particle physics*, in *Proceedings of the 1974 CERN School of Computing*. 1974. https://cds.cern.ch/record/186223.

[112] CMS Collaboration,, Phys. Lett. **B729** (2014) 149–171, arXiv:1311.7667 [hep-ex].

[113] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, and C. Wanotayaroj, Journal of High Energy Physics **2015** no. 2, (2015) 75, http://dx.doi.org/10.1007/JHEP02(2015)075.

[114] ATLAS Collaboration Collaboration,, *Jet Re-clustering in ATLAS*, Tech. Rep. ATL-COM-PHYS-2014-1117, CERN, Geneva, Sep, 2014. https://cds.cern.ch/record/1753401.

[115] D. Krohn, J. Thaler, and L.-T. Wang, JHEP **06** (2009) 059, arXiv:0903.0392 [hep-ph].

[116] ATLAS Collaboration, G. Aad et al., JHEP **09** (2013) 076, arXiv:1306.4945 [hep-ex].

[117] J. Kearney, A. Pierce, and J. Thaler, JHEP **08** (2013) 130, arXiv:1304.4233 [hep-ph].

[118] ATLAS Collaboration Collaboration,, *Identification of Hadronically-Decaying W Bosons and Top Quarks Using High-Level Features as Input to Boosted Decision Trees and Deep Neural Networks in ATLAS at $\sqrt{s}$ = 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2017-004, CERN, Geneva, Apr, 2017. https://cds.cern.ch/record/2259646.

[119] ATLAS Collaboration, G. Aad et al., Eur. Phys. J. **C76** no. 5, (2016) 295, arXiv:1512.05955 [hep-ex].

[120] J. Cogan, M. Kagan, E. Strauss, and A. Schwarztman, JHEP **02** (2015) 118, arXiv:1407.5675 [hep-ph].

[121] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, JHEP **07** (2016) 069, arXiv:1511.05190 [hep-ph].

[122] J. Schmidhuber, CoRR **abs/1404.7828** (2014), http://arxiv.org/abs/1404.7828.

[123] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[124] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer Series in Statistics. Springer New York, 2009.

[125] J. S. Conway, R. Bhaskar, R. D. Erbacher, and J. Pilot, Phys. Rev. **D94** no. 9, (2016) 094027, arXiv:1606.06859 [hep-ex].

[126] F. Chollet, *keras*, https://github.com/fchollet/keras, 2017.

[127] Theano Development Team, arXiv e-prints **abs/1605.02688** (2016), http://arxiv.org/abs/1605.02688.

[128] D. P. Kingma and J. Ba, CoRR **abs/1412.6980** (2014), http://arxiv.org/abs/1412.6980.